

THESIS / THÈSE

MASTER EN SCIENCES INFORMATIQUES

Modélisation et résolution d'un problème majeur de production d'acier à partir de la théorie des ensembles flous

Piret, Fabian

Award date:
1995

Awarding institution:
Université de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Mémoire en vue de l'obtention du grade de Maître en Informatique

**Modélisation et
résolution d'un problème
majeur de production
d'acier à partir de la
théorie des ensembles
flous**

Promoteur externe : Monsieur Marc Lescrenier (profilARBED)

Promoteur interne : Docteur Jean-Pol Leclercq

Fabian Piret

Facultés Universitaires Notre-Dame de la Paix de Namur

Année académique 1995-1996

BS 6850269

307367

Je remercie mes parents grâce à qui j'ai pu réaliser ces études et sans qui je n'aurais pu obtenir cet instrument de travail qui m'a permis de rédiger dans de bonnes conditions.

Je remercie Marc Luppi et Didier Gohy pour m'avoir relu et pour la patience qu'il ont eue en m'écoutant épiloguer sur ce travail.

Je remercie Jacques Gustin pour les 22 heures de lecture et les 700 commentaires souvent justifiés qu'il m'aura laissés.

Je remercie Monsieur Leclercq pour ses conseils judicieux et les conversations intéressantes sur ce mémoire.

Enfin, je remercie particulièrement Marc Lescrenier qui a toujours eu à cœur d'améliorer ce travail. Notre collaboration fut brève mais fructueuse. Les frontières de mon esprit critique et de mon expérience professionnelle ne pourrons que mieux s'en porter dans l'avenir.

Abstract

Since a few years, the steel industrial production knows a revival based on the old iron recycling. In collaboration with the profilARBED society, we try to formalize and to solve one production problem of great importance. The grapeshots composition characteristics are known approximatively. We have to find it mathematically.

In this aim, we develop and we test a solution prototype based on a multiobjective fuzzy linear regression method. The fuzzy sets theory, that we introduce briefly, is in fact used to treat problems where datas are imprecise. The results obtained are encouraging. We hope this work will permit the profilARBED engineers to improve this method and to exploit it at an industrial level.

Depuis quelques années, la production industrielle d'acier connaît un renouveau basé sur le recyclage de ferrailles. En collaboration avec la société profilARBED, nous essayons de formaliser et de résoudre un problème majeur de production. La composition des mitrilles est connue approximativement. Nous devons la retrouver mathématiquement.

C'est dans cet objectif que nous développons et testons un prototype de solution basé sur une méthode de régression linéaire multiobjectifs floue. La théorie des ensembles flous, que nous introduisons brièvement, est en fait utilisée pour traiter des problèmes où les données sont imprécises. Le résultat obtenu est encourageant. Nous espérons que ce travail permettra aux ingénieurs de profilARBED d'améliorer cette méthode et de l'exploiter à un niveau industriel.

Je remercie mes parents grâce à qui j'ai pu réaliser ces études et sans qui je n'aurais pu obtenir cet instrument de travail qui m'a permis de rédiger dans de bonnes conditions.

Je remercie Marc Luppi et Didier Gohy pour m'avoir relu et pour la patience qu'il ont eue en m'écoutant épiloguer sur ce travail.

Je remercie Jacques Gustin pour les 22 heures de lecture et les 700 commentaires souvent justifiés qu'il m'aura laissés.

Je remercie Monsieur Leclercq pour ses conseils judicieux et les conversations intéressantes sur ce mémoire.

Enfin, je remercie particulièrement Marc Lescrenier qui a toujours eu à cœur d'améliorer ce travail. Notre collaboration fut brève mais fructueuse. Les frontières de mon esprit critique et de mon expérience professionnelle ne pourrons que mieux s'en porter dans l'avenir.

Principales abréviations et notation

Les définitions et les paramètres inconnus sont écrits en gras.
Les parties de phrase soulignées sont considérées comme des points clés.

SC	sous contrainte(s)
TQ	tel(s) que - telle(s) que
ssi	si et seulement si
PL	programme linéaire

Table des matières

1. INTRODUCTION	5
1.1 OBJECTIFS DU STAGE ET DU MEMOIRE	5
1.2 PRESENTATION DE L'ARBED	5
1.2.1 Généralités	5
1.2.2 La société profilARBED	6
1.2.3 Les nouvelles filières de production du secteur des produits longs.....	6
1.2.3.1 Un schéma simplifié.....	6
1.2.3.2 Le jargon de la sidérurgie.....	8
1.2.3.3 Trois problèmes majeures de production.....	9
2. MODELISATION DES TROIS PROBLEMES DE PRODUCTION	11
2.1 CONVENTIONS DE MODELISATION.....	11
2.1.1 Ordre de grandeur.....	12
2.1.2 Modélisation du circuit de production	13
2.2 MODELISATION DU PROBLEME DIRECT	13
2.2.1 Hypothèses de production.....	13
2.2.2 Enoncé formel.....	14
2.2.3 Solution actuelle.....	15
2.2.4 Insuffisances de la solution actuelle.....	16
2.3 MODELISATION DU PROBLEME INVERSE	17
2.3.1 Conventions supplémentaires de notation	17
2.3.2 Hypothèses de production.....	17
2.3.3 Enoncé formel dans le cas 1	18
2.3.4 Enoncé formel dans le cas 2	19
2.3.5 Solution actuelle dans le cas 2.....	19
2.3.6 Insuffisances de la solution actuelle.....	20
2.4 LE PROBLEME DE PLANIFICATION DE PRODUCTION	21
3. INTRODUCTION A LA THEORIE DES ENSEMBLES FLOUS.....	23
3.1 GENERALITES	23
3.1.1 Quelques premiers commentaires	24
3.2 LA THEORIE DES PROBABILITES ET LA THEORIE DES POSSIBILITES	25
3.2.1 Un mot sur l'algèbre des événements.....	25
3.2.2 Les fondements de la théorie des probabilités.....	27
3.2.2.1 Le concept intellectuel sous-jacent.....	27
3.2.2.2 Les limites pratiques de la théorie des probabilités.....	28
3.2.2.3 Les axiomes et définitions de base dans le cas discret	28
3.2.2.4 Dualité mesure de probabilité-distribution de probabilité.....	29
3.2.2.5 L'évaluation d'une mesure de probabilité	29
3.2.2.6 La suite de la théorie des probabilités.....	30
3.2.3 Les fondements de la théorie des possibilités	30
3.2.3.1 Pourquoi la théorie des possibilités ?	30
3.2.3.2 Le concept intellectuel sous-jacent.....	31
3.2.3.3 Les axiomes et définitions de base dans le cas discret	32
3.2.3.4 La dualité mesure de possibilité-distribution de possibilité	33
3.2.3.5 L'évaluation d'une mesure de possibilité	33
3.2.3.6 Mesure de nécessité.....	34
3.2.3.7 La suite de la théorie des possibilités	34
3.2.4 Une comparaison des deux théories.....	35
3.3 LA THEORIE DES ENSEMBLES FLOUS.....	35
3.3.1 Le concept intellectuel sous-jacent.....	35
3.3.2 Les définitions de base.....	37
3.3.2.1 Les sous-ensembles flous.....	37
3.3.2.2 Les α -coupes, le support et le support topologique d'un sous-ensemble flou	39
3.3.2.3 Les nombres flous	40
3.3.2.4 Le principe d'extension en général	43
3.4 LA THEORIE POSSIBILISTE DES ENSEMBLES FLOUS	44
3.4.1 La justification possibiliste du principe d'extension de Zadeh, Dubois et Prade.....	44

3.4.2 Le principe d'extension proposé par Zadeh, Dubois et Prade	45
3.4.3 Une propriété intéressante de Hung T. Nguyen	47
3.4.4 Trois indices de comparaison de nombres flous proposés par Dubois et Prade	47
3.4.5 Le calcul de la multiplication et de l'addition de nombres flous	48
4. A LA RECHERCHE D'UNE SOLUTION « FLOUE » A UN DES TROIS PROBLEMES	51
4.1 QUE CHERCHER	51
4.1.1 ... pour le problème direct ?	51
4.1.1.1 Les hypothèses	51
4.1.1.2 Un énoncé flou du problème direct	52
4.1.2 ... pour le problème inverse ?	52
4.1.2.1 Conventions supplémentaires de notation	52
4.1.2.2 Les hypothèses	53
4.1.2.3 Un énoncé flou du problème inverse	53
4.1.3 ... pour le problème de planification de production ?	54
4.2 OU CHERCHER ?	54
4.3 LES RESULTATS DES PREMIERES RECHERCHES	55
4.3.1 Sur InterNet	55
4.3.2 Par courrier et fax	56
4.3.3 Dans les revues et livres spécialisés	56
4.4 LA SELECTION D'UN DES TROIS PROBLEMES	57
5. LE MODELE SAKAWA-YANO COMME SOLUTION AU PROBLEME INVERSE.	59
5.1 PRESENTATION DU MODELE	59
5.2 LES HYPOTHESES DU MODELE	59
5.2.1 Les correspondances de notation	59
5.2.2 Les hypothèses mathématiques	60
5.3 SATISFIABILITE DES HYPOTHESES SAKAWA-YANO	60
5.3.1 Les variables d'entrée	61
5.3.2 Les variables de sortie	61
5.3.3 Les variables indépendantes	61
5.4 LES PRINCIPALES SIMPLIFICATIONS DU MODELE SAKAWA-YANO	62
5.4.1 Les partitions $J1$, $J2$ et $J3$ sont connues	62
5.4.2 L'expression des α -coupes est plus simple	62
5.4.3 L'expression de la fonction objectif commune	62
5.4.4 Les deux premiers problèmes sont les mêmes	63
5.4.5 Le problème $P3(\alpha)$ n'admet sémantiquement pas de solution	63
5.5 APPLICATION DE LA SOLUTION SAKAWA-YANO	64
5.6 LES PREMIERES CRITIQUES DE LA SOLUTION SAKAWA-YANO	67
6. TESTS DE LA SOLUTION SAKAWA-YANO	69
6.1 COMMENT TESTER LE MODELE ?	69
6.2 LE GENERATEUR DE TEST	71
6.2.1 Les paramètres de simulation	71
6.2.1.1 Les paramètres globaux	71
6.2.1.2 Les paramètres du comportement de variation des teneurs	71
6.2.1.3 Les paramètres du comportement des erreurs commises sur les mesures	71
6.2.1.4 La fonction $L(.)$ et la dispersion sur les nombres flous	72
6.2.2 La politique de génération virtuelle des charges	72
6.3 LA PRESENTATION DES RESULTATS	73
6.3.1 Notre méthodologie de test	73
6.3.1.1 Le tableau 1	74
6.3.1.2 Le tableau 2	74
6.3.2 Les tests de type 1	75
6.3.3 Les tests de type 2	76
6.3.4 Les tests de type 3	80
6.3.5 Les tests de type 4	85
6.4 LA COMPREHENSION DES RESULTATS	85
6.4.1 Les tests de type 1	86
6.4.2 Les tests de type 2	86
6.4.3 Les tests de type 3 et 4	88
6.5 LA SOLUTION FINALE PROPOSEE	88

7. CONCLUSION GENERALE.....	91
8. BIBLIOGRAPHIE.....	93
8.1 INTRODUCTION A LA THEORIE DES ENSEMBLES FLOUS	93
8.2 LE PROBLEME DIRECT.....	93
8.3 LE PROBLEME INVERSE.....	93
8.4 LE PROBLEME DE PLANIFICATION DE PRODUCTION	94

Chapitre 1

Introduction

1. Introduction

1.1 Objectifs du stage et du mémoire

Notre stage s'est déroulé au sein d'une aciérie de grand renom, l'ARBED. Les objectifs étaient :

- Etudier cette « nouvelle » branche très à la mode des mathématiques qu'est la théorie des ensembles flous (fuzzy sets). Il semble qu'elle soit bien adaptée à la résolution de problèmes industriels, c.-à-d. dans lesquels les données manipulées sont souvent imprécises.
- Comprendre trois problèmes majeurs des nouvelles filières de production d'acier, et étudier les potentialités du flou pour une « meilleure » résolution de chacun d'eux. Ces trois problèmes sont respectivement, le problème direct, le problème inverse, et la planification de production.
- Sélectionner un des trois problèmes, avec une préférence pour le problème inverse, au vu de ces potentialités. Pour ce problème, implémenter un prototype d'application industrielle basé sur le flou.
- Tester ce prototype, et éventuellement passer à une application industrielle si les résultats des tests sont satisfaisants.

1.2 Présentation de l'ARBED

1.2.1 Généralités

Le groupe international ARBED est incontestablement un leader mondial dans le secteur de la production de produits finis ou semi-finis à base d'acier. Avec une production de 11,5 millions de tonnes d'acier, le groupe figure parmi les premières sociétés sidérurgiques du monde. Il compte actuellement pour un tiers de l'économie luxembourgeoise. La maison-mère se trouve à Luxembourg-ville. Elle fut construite en 1882 et représente une magnifique imitation d'architecture antique. Le groupe comporte une dizaine de centres de production et une quarantaine de réseaux de vente. Approximativement la moitié de ceux-ci se situent en Europe, terre natale de l'entreprise. Ses activités se répartissent sur les huit secteurs industriels suivants :

- le secteur des produits plats
- le secteur des produits longs
- le secteur des aciers inoxydables
- le secteur commercial
- le secteur des tréfileries
- le secteur de l'engineering et des constructions mécaniques
- le secteur des feuilles de cuivre
- le secteur Brésil

1.2.2 La société profilARBED

Dans le secteur industriel des produits longs, c'est la société profilARBED qui contrôle l'ensemble des activités. Elle fut fondée en novembre 1992. Ce secteur comprend les sites de production d'Esch-Belval et de Differdange (Luxembourg) opérant principalement dans la gamme des poutrelles et des profilés moyens et lourds.

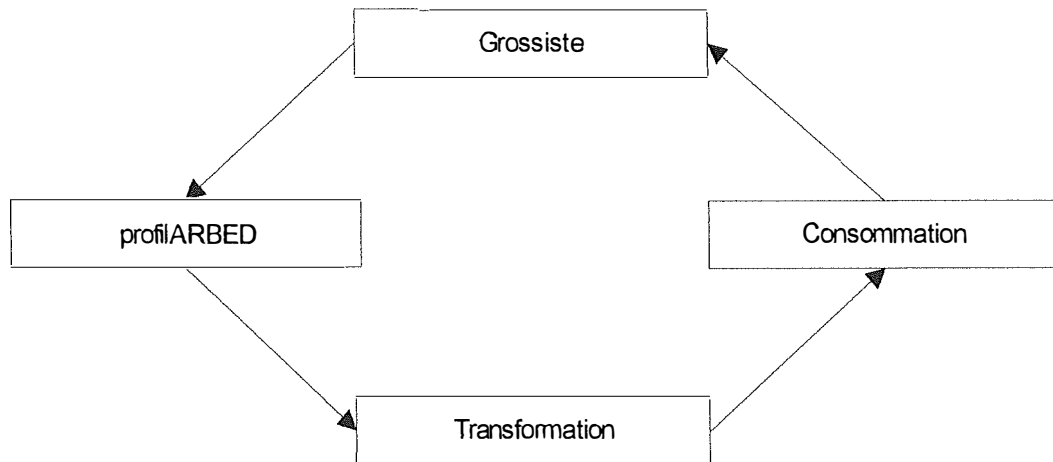
Comme toute entreprise de pointe, profilARBED possède un ensemble de services informatiques. Sa structure hiérarchique se découpe en Services, eux-même divisés en Circuits, à nouveau divisés en Unités. Notre stage se déroule au sein de « l'Unité Modélisations et Optimisations » du « Circuit Production » au « Service Informatique ». Cette unité est notamment responsable de projets traitant de la modélisation et de l'optimisation des procédés à mettre en œuvre pour la gestion des nouvelles filières de production.

1.2.3 Les nouvelles filières de production du secteur des produits longs

1.2.3.1 Un schéma simplifié

Les anciennes filières de production sont basées sur la technologie des hauts-fourneaux. A partir de différents minerais, on produit de l'acier. Aujourd'hui, ces filières sont devenues obsolètes. Depuis quelques années, elles sont remplacées par de nouvelles, plus rentables, basées sur le recyclage de ferrailles. A titre d'ordre de grandeur, la durée d'un cycle de production était approximativement de 10h dans les anciennes filières. Il est de 1h dans les nouvelles ... L'idée est la suivante. On récupère de la mitraille avec laquelle on produit de l'acier qui, après être passé dans les mains du consommateur, finit par revenir sous forme de mitraille dans les mains de grossistes ! On obtient ainsi des filières de production cycliques, dont la rentabilité est accrue par des techniques de production nouvelles telles que les fours électriques ou la coulée continue. En ce qui concerne profilARBED, le processus de glissement du secteur des produits longs vers les filières électriques est entamé depuis 1992 et se terminera en 1996. Il est terminé sur le site de Differdange et est dans ses dernières phases sur le site d'Esch-Belval.

Cette filière de production cyclique peut être symboliquement découpée en quatre phases que nous appelons respectivement la phase Grossiste, la phase profilARBED, la phase de Transformation, et la phase de Consommation. Ce schéma symbolise leur enchaînement.



Les activités de profilARBED se situent principalement au niveau de la phase portant son nom. Elles s'étendent aussi en amont, dans la phase Grossiste, lors de l'acheminement des mitrailles à l'usine. C'est ainsi que l'on peut régulièrement apercevoir sur les voies de chemin de fer des wagons portant le sigle ARBED. Les activités de profilARBED recouvrent aussi la phase de Transformation, lors de la production de certains produits finis à base d'acier.

La phase grossiste

Des grossistes récupèrent de la ferraille dans la nature. Ils les regroupent selon leur composition. Les entreprises telles que profilARBED font partie de leurs clients.

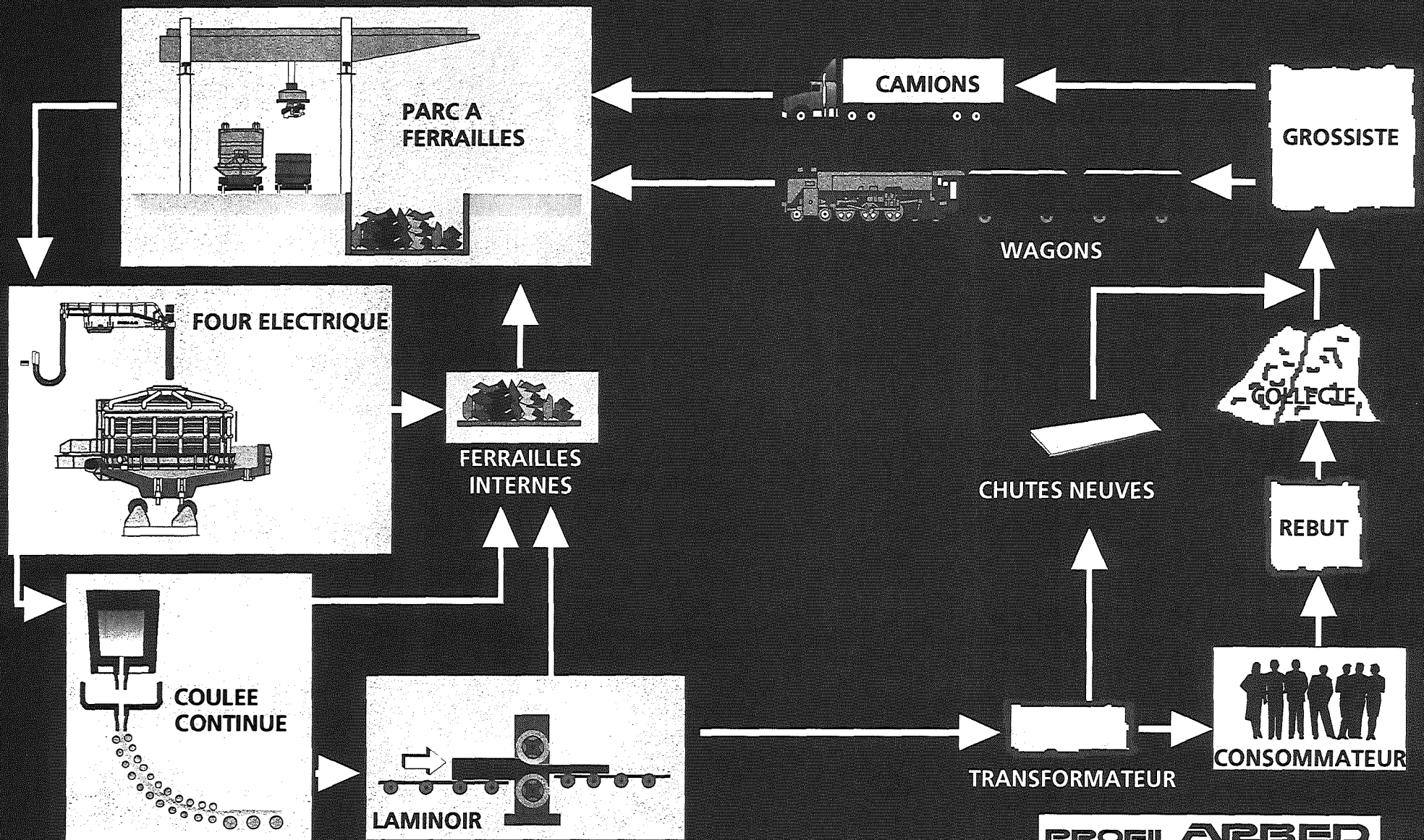
La phase profilARBED

La société profilARBED achète ces différentes ferrailles à ces grossistes (fournisseurs). Ces ferrailles sont transformées en mitraille. Le prix des mitrailles se discutera en fonction de leurs qualités. Les plus prisées, et donc celles que nous considérerons de bonne qualité (c.à.d. particulièrement désignées pour le chargement des fours électriques en continu) doivent avoir une teneur en fer élevée (plus de 95%), une teneur en impuretés métalliques résiduelles faible, et être denses et homogènes. Ainsi, la société profilARBED sous-classifie ces mitrailles en différents types de mitraille, afin de pouvoir distinguer finement leurs origines. Chaque type de mitraille est identifié par un triplet composé de son fournisseur (aux environs de 50), de sa catégorie (14 catégories possibles : mitraille de collecte, de chute d'atelier,...), et de son chantier de provenance (3 chantiers possibles). Parmi les 2100 ($50 \times 14 \times 3$) combinaisons possibles, 1000 existent réellement et correspondent à 1000 types de mitrailles utilisés. Notons que ces 1000 types de mitraille ne sont jamais physiquement disponibles en même temps, tout au long de l'année. Tout au plus 50 sont utilisés simultanément pendant un certain temps. Les 1000 types de mitraille sont tous utilisés par un roulement des fournisseurs tout au long de l'année. A partir de ces types de mitraille, profilARBED produit de l'acier sous forme de différents produits semi-finis tels que des poutrelles métalliques.

La phase de transformation

Certaines entreprises achètent ces produits semi-finis à profilARBED pour en faire des produits finis tels que des ponts ou des bâtiments. Lors de leur production il leur reste

Flux des ferrailles



PROFIL ARBES

évidemment des chutes. Celles-ci retournent directement chez les grossistes, pour constituer une mitraille de bonne qualité. Le reste est donné en pâture à la société de consommation.

La phase de consommation

Ces produits finis utilisés par la société de consommation terminent sous forme de rebut. Le temps de retour d'une ferraille, soit le temps entre le moment où elle sort de profilARBED et le moment où elle termine sous forme de mitraille, peut varier de plusieurs semaines, dans le cas de boîtes de conserve, à plusieurs dizaines d'années dans le cas de bâtiments. Techniquement, il serait possible de récupérer 70% des métaux ferreux. C'est ce que prétend déjà faire le Japon, ce dont il convient raisonnablement de douter. Historiquement, ce taux ne semble guère dépasser 35 à 40%.

1.2.3.2 *Le jargon de la sidérurgie*

Voici un schéma plus détaillé de ces nouvelles filières de production cycliques. Il permet de mettre en évidence certains termes du jargon sidérurgique. Ces termes, inscrits en gras, seront constamment utilisés tout au long de la présentation. Il est important de bien les définir afin de bien les distinguer.

L'acier est un alliage de différents métaux, vérifiant certaines contraintes physiques (densité, élasticité, résistance mécanique, ...) et chimiques (pourcentages admissibles des différents métaux, ...). Le métal le plus présent est de loin le fer (en général plus de 95%). D'autres métaux tels que le cuivre (moins de 0.5%) et le nickel (moins de 0.1%) sont aussi présents. Nous appelons l'ensemble de ces métaux, **les éléments chimiques** de l'acier. Par abus de langage nous les appellerons aussi **éléments**. Nous en comptons 21 au total.

La partie de production qui nous intéresse est située entre le parc à ferrailles et le laminoir. Dans le parc à ferraille, profilARBED dispose d'un certain nombre de **types de mitraille**. Comme nous l'avons déjà expliqué, chaque type de mitraille correspond à une mitraille dont on connaît plus ou moins finement les origines. La composition de ces types de mitraille est donnée par les fournisseurs, et constitue ce que l'on appelle **la carte d'identité des mitrailles**. Un type de mitraille est constitué de différents éléments chimiques, dont ceux nécessaires à la production d'acier. On peut y retrouver d'autres composants comme de l'oxyde de carbone, de l'oxyde de fer, du caoutchouc, du plastique, de l'eau, ou encore du bois. Le but de ce circuit de production est de produire un certain nombre de profilés lourds d'acier à partir des différents types de mitraille. Un profilé lourd d'acier est une sorte de grande poutrelle d'acier. En sortie du laminoir, celle-ci est découpée en petits morceaux appelés lingots, un produit semi-fini. A partir de ces lingots, on peut produire des articles en « inox » ou des ponts. La composition de l'acier, de chaque profilé lourd, est homogène et déterminée à l'avance, pour chaque élément.

Nous appelons **une charge**, un mélange de différentes quantités de chaque type de mitraille. Chaque charge peut se présenter sous différentes « formes » physiques et chimiques. L'évolution de sa forme initiale, en entrée du circuit de production, à sa forme finale, en sortie du circuit de production, est constituée d'une série d'étapes intermédiaires (moins nombreuses que dans les anciennes filières de production). La durée de production est approximativement de 1 heure. En sortie du circuit de production, cette charge sous d'un forme d'un profilé lourd d'acier. En entrée du circuit de production, la « forme » d'une charge est un ensemble de tas, chacun pris dans un type de mitraille. Le tout pèse aux

environs des 110 tonnes, et est enfourné dans le four électrique, pour terminer en forme de poutrelle d'acier d'à peu près 100 tonnes. Les 10 tonnes d'écart représentent l'ensemble des composants qu'il faut éliminer. L'ensemble des étapes intermédiaires sert simplement à fondre la charge sous « sa forme initiale », à éliminer les 10 tonnes de composants inutiles (bois, plastique...), à ne garder que les bonnes proportions de chaque élément chimique de l'acier, et à transformer le tout en une poutrelle, « la forme finale » de la charge. Cela paraît simple, mais les processus chimiques sous-jacents sont relativement compliqués, notamment en ce qui concerne l'élimination du cuivre.

Chaque profilé lourd produit est systématiquement analysé. C'est ainsi que l'on détermine très précisément la teneur de l'acier en chacun de ses éléments. Nous appelons cette étape, l'analyse de la charge. Grâce à cette analyse, il est possible de vérifier que la qualité de l'acier est bien identique à celle que l'on attendait.

1.2.3.3 Trois problèmes majeures de production

Après cette présentation grossière de l'aspect technique des nouvelles filières de production d'acier, voici quelques mots sur leur exploitation proprement dite. La rentabilité de tels circuits de production passe par la résolution d'un certain nombre de **problèmes de production**. En voici trois parmi les plus importants.

Le premier est appelé **problème direct**. On désire produire un certain nombre de charges de composition connue en chacun des éléments chimiques. C'est l'analyse désirée des charges. Pour chacune des charges, le problème consiste à déterminer la quantité de chaque type de mitraille à enfourner pour obtenir cette analyse désirée. L'analyse mesurée des charges, en sortie du circuit de production, est ensuite comparée à l'analyse désirée. Idéalement l'analyse mesurée et l'analyse désirée sont les mêmes.

Le deuxième est appelé **problème inverse**. A cause de certaines erreurs sur les données de composition des différents types de mitraille, les charges produites à partir de la résolution du problème direct, ont souvent une composition différente de celle attendue. En réalité, la qualité des types de mitraille (annoncée dans leur carte d'identité par les fournisseurs) est souvent erronée. Par exemple, les types de mitrilles ne sont pas homogènes, ou des fournisseurs exagèrent la qualité de leur mitraille. Idéalement, il serait donc utile de localiser de façon sûre les mitrilles les moins homogènes indépendamment de ce qu'annoncent les fournisseurs. On peut même essayer de retrouver la composition exacte de chaque type de mitraille. Remarquons que cet objectif est plus général que le précédent. Il ne se contente pas seulement de repérer les « vilains » fournisseurs, mais retrouve aussi les proportions dans lesquelles ils se trompent. C'est là, le but du problème inverse. Inverse, simplement parce que l'on tente de retrouver des données utiles, en l'occurrence les teneurs des mitrilles, en entrée du circuit de production, à partir des analyses mesurées en sortie du circuit de production. Ce problème est intéressant à résoudre à plusieurs égards. D'une part, il permettrait de disposer de données plus correctes concernant la composition des mitrilles. Il augmenterait donc les chances que les analyses mesurées des charges produites soient égales à celles attendues. D'autre part, il permettrait de repérer les types de mitraille de moins bonne qualité, et ainsi permettrait de renégocier leur prix auprès de leur fournisseur. Le gain financier à réaliser ainsi est loin d'être négligeable, étant donné les énormes quantités de mitraille achetées et l'amélioration de la rentabilité de production qu'il en découle.

Les problèmes inverse et direct sont donc complémentaires. Le problème inverse existe, pour corriger les erreurs qui ne permettent pas au problème direct de remplir son rôle de manière satisfaisante. ProfilARBED utilise une solution informatique pour résoudre le problème direct. Le problème inverse, lui, est « résolu » empiriquement, grâce à l'expérience et l'intelligence des ingénieurs de production. Cette manière de faire n'est pas satisfaisante. En effet, les ingénieurs n'arrivent qu'à repérer intuitivement les fournisseurs qui se trompent dans de grosses proportions, mais pas les autres. De plus, pour pouvoir incriminer un fournisseur, des arguments intuitifs ne suffisent pas ! ProfilARBED a bien tenté de développer une solution informatique, mais comme nous le verrons, celle-ci ne donnent pas de résultats satisfaisants.

Le troisième et dernier problème que nous abordons est appelé le **problème de planification de production**. Alors que les deux premiers traitent du « quoi », celui-ci traite du « quand » ! En effet, il a pour but de déterminer à quel moment il faut enfourner les différentes quantités de chaque type de mitraille d'une charge sous sa forme initiale. Il serait naïf de croire que l'enfournement de la charge $j+1$ se fait lorsque la charge j est produite. Etant donné que la chaîne de production est composée de différentes étapes, pourquoi ne pas déjà enfourner (première étape) la charge $j+1$ lorsque la charge j passe à la coulée continue (deuxième étape). A un moment donné, il y a donc plusieurs charges en cours de production, toutes à des étapes différentes. La gestion de ce parallélisme de production est la base de ce problème.

Chapitre 2

Modélisation des trois problèmes de production

2. Modélisation des trois problèmes de production

Pour chacun de ces trois problèmes de production, nous allons maintenant tenter de trouver une modélisation mathématique correcte. Nous présentons donc trois modèles différents et indépendants. Chacun d'eux est constitué de quatre grandes parties : des hypothèses, un énoncé, une solution, et les insuffisances de cette solution. Les hypothèses d'un modèle servent à déterminer, d'une part, le domaine de valeurs de chacun des paramètres du modèle, et d'autre part, lesquels parmi ceux-ci sont inconnus. Lorsqu'ils sont connus, nous expliquons la façon dont ils sont déterminés en production . Pour chaque modèle, il est important de bien en comprendre les hypothèses. L'énoncé d'un modèle est une formulation plus ou moins formelle du problème correspondant en terme des différents paramètres. Cette étape est nécessaire, car jusqu'à présent, l'énoncé des problèmes est encore fort ambigu. Chaque problème, est ainsi traduit sous forme d'un énoncé mathématique, pour lequel on ne spécifie pas encore (exactement) comment on essaye d'en déterminer les inconnues. Il s'agit du quoi, des spécifications du problème. La solution d'un modèle, représente une solution mathématique (choisie par profilARBED jusque maintenant) à l'énoncé du problème. Nous expliquons comment on détermine mathématiquement les inconnues du problèmes, en répondant aux spécifications de son énoncé. Enfin, les insuffisances de la solution, mettent en évidence les raisons pour lesquelles la solution choisie ne répond pas aux spécifications de son énoncé.

Modèle d'un problème de production
Hypothèses de production
Enoncé formel
Solution actuelle
Insuffisances de la solution

Tous ces modèles utilisent un certains nombre de notations et de conventions communes. Avant de passer à leur modélisation respective, nous présentons ces conventions. Nous nous permettons de souligner leur importance pour la suite.

Bien que les hypothèses de chacun des modèles utilisent des paramètres identiques (en notation), il est important de signaler que leur domaine de valeur et leur statut (en tant qu'inconnue ou pas) varient selon les modèles. Ceci semblera peut-être difficile à gérer, étant donnée la forte dualité entre le problème direct et le problème inverse. Ceci est du au fait que les insuffisances de la solution au problème direct, influent sur les hypothèses du problème inverse, et vice-versa. Nous aurions pu utiliser des notations différentes, pour chacun d'eux. Nous avons décidé de ne pas choisir cette optique, car il s'ensuivrait une explosion du nombre de notations

2.1 Conventions de modélisation

Pour tenter de résoudre les problèmes présentés ci-dessus, nous utiliserons des modèles « mathématico-physico-chimiques ». Ils auront tous en commun les notations suivantes :

$c \in \mathcal{C} \equiv$	l'ensemble des indices de charges	$\subseteq \mathbb{N}$	avec	$C = \#\mathcal{C}$
$m \in \mathcal{M} \equiv$	l'ensemble des indices de types de mitraille	$\subseteq \mathbb{N}$	avec	$M = \#\mathcal{M}$
$e \in \mathcal{E} \equiv$	l'ensemble des indices d'élément chimique	$\subseteq \mathbb{N}$	avec	$E = \#\mathcal{E}$
$t(m,e) =$	teneur de la mitraille m en l'élément e (« t » pour teneur)			
$q(c,m) =$	poids de la quantité du type de mitraille m utilisée pour la production de la charge c (« q » pour quantité)			
$a(e,c) =$	poids de l'élément e dans la charge c (« a » pour analyse)			

Nous pouvons exprimer ces mêmes données sous des formes matricielles redondantes, qui nous arrangeront par la suite :

Q	=	matrice ($C \times M$) dites des quantités
A_e	=	vecteur à (C composantes) dit des analyses en l'élément e
T_e	=	vecteur (M composante) dit des teneurs en l'élément e
définis comme suit :		
$[Q]_{m,c} = q(c,m)$	$\forall c \in \mathcal{C}, \forall m \in \mathcal{M}$	
$[A_e]_c = a(e,c)$	$\forall e \in \mathcal{E}, \forall c \in \mathcal{C}$	
$[T_e]_m = t(m,e)$	$\forall e \in \mathcal{E}, \forall m \in \mathcal{M}$	

Nous ne déterminons pas les domaines de valeur et le statut (en tant qu'inconnue ou pas) des quantités $q(.,.)$, $t(.,.)$, et $a(.,.)$. Comme nous l'avons déjà signalé, ils changeront au fur et à mesure des hypothèses que nous ferons dans nos modélisations des trois problèmes. Par exemple nous verrons que $a(e,c)$ est un paramètre connu dans les problèmes direct et inverse. Cependant, dans le premier problème, il s'agit d'une valeur désirée avant la production, alors que dans le deuxième c'est une valeur mesurée après la production. Dans l'état actuel des choses, ces deux valeurs sont souvent différentes, mais nous gardons la même notation pour les deux. Plus brièvement, en terme de notation, nous ne ferons pas la différence entre un paramètre désiré, et un paramètre mesuré.

2.1.1 Ordre de grandeur

Il est à notre avis intéressant de fournir un ordre de grandeur des différents paramètres renseignés ci-dessus. Ceci vous permettra d'imaginer la taille des problèmes gérés en production à profilARBED.

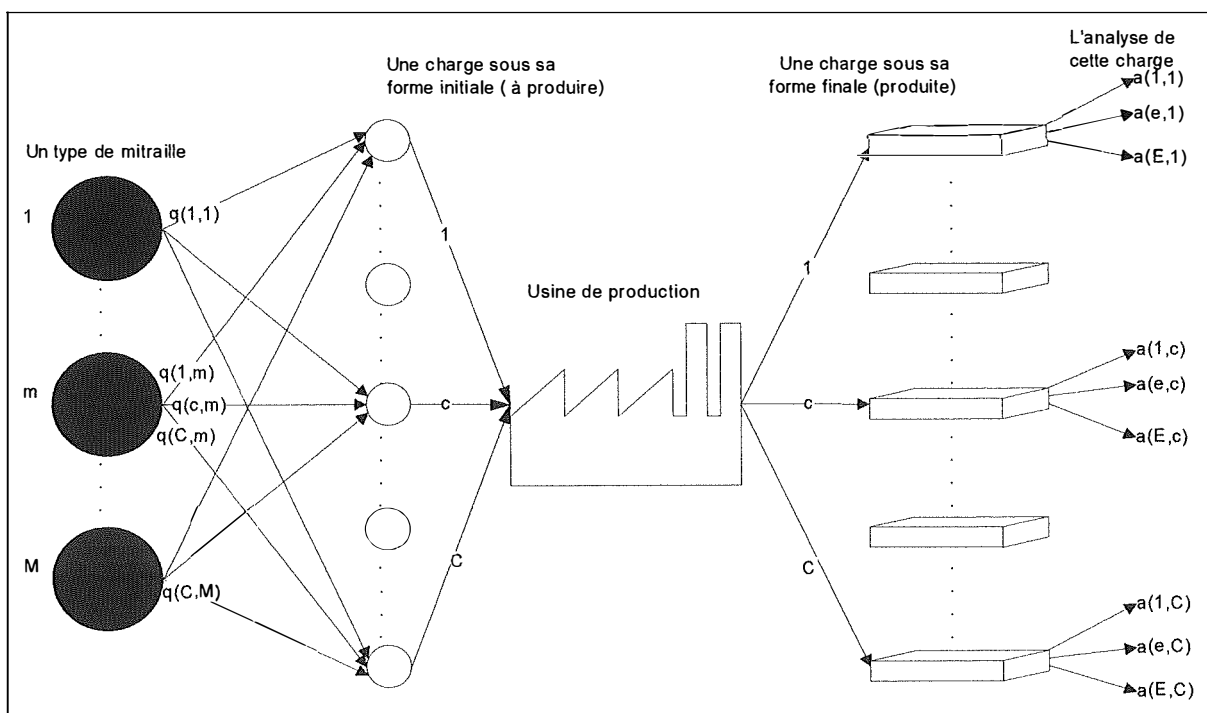
Paramètre	Ordre de grandeur	Raison
E	21	Il y a 21 éléments répertoriés dans l'acier. Tous ne sont pas toujours présents en fonction de la qualité de l'acier
M	50	Il y en a en réalité ± 1000 , mais seulement 50 sont disponibles à la fois
C	200	C'est le nombre maximal de charge que l'on estime pouvoir produire avec les 10 types de mitraille.

$Q(.,.)$	20 à 80 T	Pour une charge initiale de 110 tonnes, certains types de mitraille sont fortement utilisés (80 tonnes), car leur rapport qualité-prix est très bon.
$a(.,.)$	0 à 98 T	Pour une charge produite de 100 tonnes, il peut y avoir plus de 95 tonnes de fer (élément principal de l'acier), jusqu'à 0 tonnes de cuivre (élément que l'on essaye d'éliminer)
$t(.,.)$	0 à 0,95	Pour des types de mitraille de bonne qualité, la teneur en fer peut être de 95% et celle en cuivre de 0%

Il est important de remarquer que C est toujours beaucoup plus grand que M . C'est en effet grâce à ceci que l'on a des chances de pouvoir résoudre correctement le problème inverse. Plus on produit de charges avec les différents types de mitraille, plus on a d'information sur la façon dont varient leurs teneurs en chacun des éléments chimiques.

2.1.2 Modélisation du circuit de production

Pour plus de clarté, voici un schéma représentant le circuit de production sur lequel nous avons greffé les différentes notation introduites ci-dessus :



2.2 Modélisation du problème direct

2.2.1 Hypothèses de production

Dans cette partie, nous déterminons le domaine de valeur et le statut (en tant qu'inconnue ou pas) de chaque paramètre du modèle, en se plaçant à la place des ingénieurs de production. Voici les hypothèses qu'ils font :

Les C charges n'ont pas encore été produites :

- Chaque type de mitraille est homogène en chaque élément.
 $\forall m \in \mathcal{M} \quad \forall e \in \mathcal{E}, t(m,e) \in \mathcal{R}$ et est connue (indiquée sur la carte d'identité du type de mitraille m)
- On connaît la qualité désirée de chaque charge à produire en chaque élément.
 $\forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}, a(e,c) \in \mathcal{R}$ et est connue. C'est l'analyse désirée.
- On cherche la quantité de chaque mitraille à utiliser pour chaque charge
 $\forall m \in \mathcal{M} \quad \forall c \in \mathcal{C}, q(c,m) \in \mathcal{R}$ et est inconnue

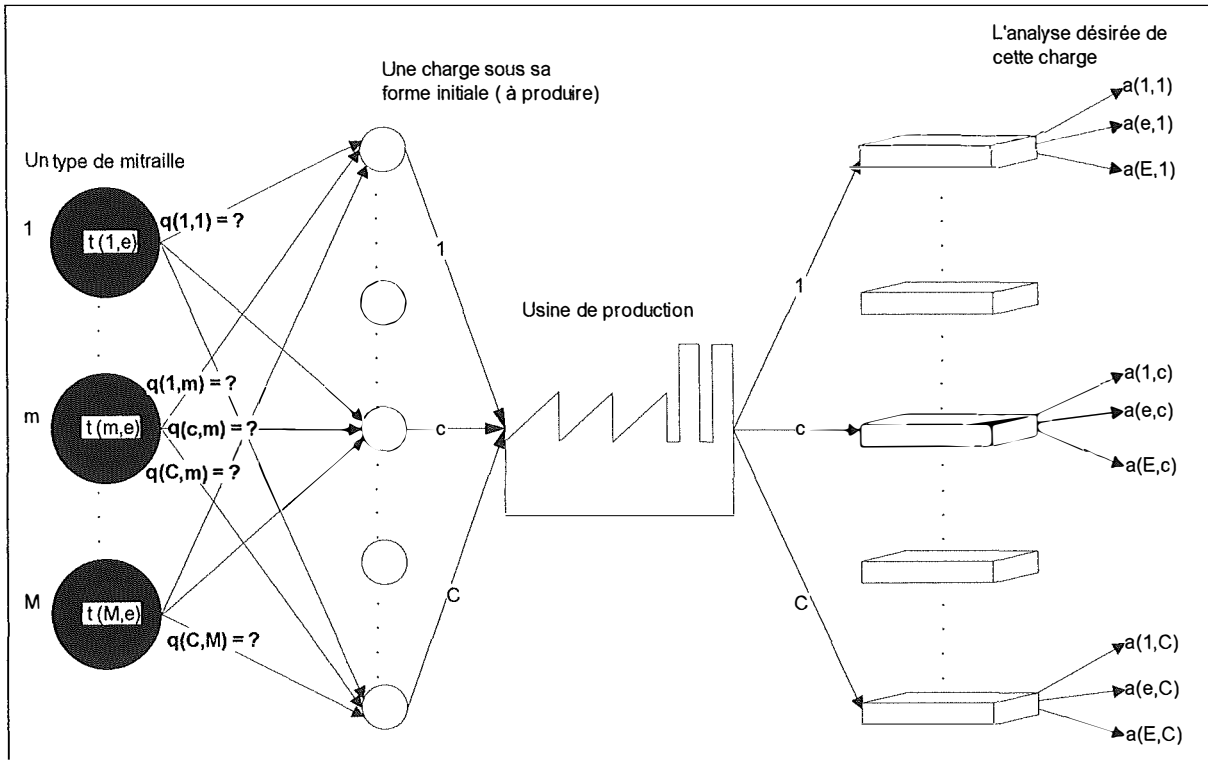
2.2.2 Enoncé formel

$\forall c \in \mathcal{C}$, choisir les « meilleures » valeurs de $q(c,m)$

$$\text{SC} \quad \sum_{m \in \mathcal{M}} t(m,e) \times q(c,m) = a(e,c) \quad \forall e \in \mathcal{E}$$

« Pour chaque charge à produire, choisir la quantité de chaque type de mitraille à enfourner, de telle manière que la composition de la charge corresponde à celle attendue ».

On représente graphiquement ce problème de la manière suivante :



2.2.3 Solution actuelle

Dans la plupart des entreprises, le critère de qualité du choix de la meilleure valeur est lié à la notion de coût. Dès lors, introduisons une nouvelle convention qui nous permet d'exprimer la notion de coût d'une charge :

$ct(m) \in \mathbb{R}$ = coût de une unité de poids du type de mitraille m

Le problème à résoudre est le suivant :

$\forall c \in \mathcal{C}$, trouver la solution du PL

$$\text{Min} \quad \sum_{m \in \mathcal{M}} ct(m) \times q(c,m)$$

$$\text{SC} \quad \sum_{m \in \mathcal{M}} t(m,e) \times q(c,m) = a(e,c) \quad \forall e \in \mathcal{E}$$

Il suffit donc de résoudre C PL à M variables et E contraintes pour résoudre le problème direct. Cette solution est implémentée informatiquement et utilisée en production.

Ceci est l'idée de base du problème. Il est évident que d'autres contraintes doivent probablement être ajoutées. Par exemple, on peut obliger le modèle à choisir les quantités de chaque type de mitraille de telle manière que leur poids total reste aux environs de 110 tonnes. En effet, si l'on n'ajoute pas ce type de contrainte, le modèle risque de choisir d'énormes quantités de types de mitraille de très mauvais qualité, car celles-ci sont très bon

marché. On risque donc de se retrouver avec des charges sous forme initiale de 250 tonnes, pour lesquelles il faudrait éliminer approximativement 150 tonnes de déchet.

2.2.4 Insuffisances de la solution actuelle

Lorsqu'on utilise cette solution pour résoudre le problème direct, un nouveau problème apparaît. Pour une charge donnée, lorsque l'on utilise les différentes quantités de chaque type de mitraille, solution de ce problème, on constate que l'analyse désirée est souvent différente de l'analyse mesurée.

Mais pourquoi ?

Si toutes les hypothèses posées pour la résolution du problème inverse étaient correctes, les analyses désirées et mesurées seraient toujours les mêmes. Comme ce n'est pas le cas, on en déduit que une ou plusieurs des hypothèses posées pour ce problème est (sont) erronée(s). Si on y regarde de plus près on comprend vite qu'il s'agit de la première, et ce pour deux raisons bien précises.

- La carte d'identité de chaque type de mitraille nous est donnée par les fournisseurs eux-mêmes. Or ceux-ci ont parfois intérêt à exagérer la qualité de leur mitraille. Par exemple, ils préfèrent annoncer que leur mitraille a une teneur en fer plus haute que la teneur réelle, ceci afin que profilARBED la leur achète à un prix plus élevé.
- En supposant que tous les fournisseurs soient honnêtes, il est physiquement impossible que leur mitraille soit homogène. Même si un type de mitraille n'est pas homogène, on peut évidemment lui associer une teneur globale. Mais ceci ne nous sert à rien car on n'enfourne jamais toute une mitraille en une fois. Pour chaque charge, on puise un petit paquet de chaque type de mitraille. C'est la forme initiale de la charge. Considérons l'ensemble des petits paquets puisés pour un type de mitraille durant la production de toutes les charges. La non homogénéité de ce type de mitraille implique que les teneurs de chacun de ces petits paquets en un élément soient (presque) toutes différentes. Dans le premier petit paquet, la teneur en fer est peut-être de 91%, alors qu'elle n'est que de 85% dans le dixième. Ceci est dû au fait que les mitrailles peuvent parfois être issues de fournisseurs ayant récupéré des matériaux de chantiers de démolition, de cimetières de voitures, ou encore de chantiers de désaffectation de bâtiments. Intrinsèquement, la composition des mitrailles est donc fort irrégulière. Les conditions de stockage et d'acheminement des mitrailles sont aussi responsables de cet irrégularité. Les mitrailles peuvent être amenées par wagons entiers et stockées dans des containers à mitraille. Ces wagons et ces containers sont évidemment utilisés plusieurs fois au cours de l'année pour différentes mitrailles. Ils sont aussi soumis aux conditions climatiques.

Mathématiquement, par rapport aux hypothèses posées pour ce problème, nous en déduisons donc que le concept de teneur d'un type de mitraille en un élément, n'a pas de sens, en tant que constante réelle pour la production de toutes les charges.

Afin d'éventuellement continuer à travailler avec des réels, nous pouvons admettre le concept de teneur d'un type de mitraille m en un élément e pour une charge c , en tant que

constante réelle. Nous noterons $\text{tr}(m,e,c)$ cette valeur réelle. Elle sera ajoutée dans les conventions de notation pour le problème inverse.

Il est aussi possible de représenter la teneur d'un type de mitraille en un élément, avec un autre concept mathématique que celui des réels. Par exemple une approche probabiliste permet de représenter la teneur d'un type de mitraille en un élément par une répartition de probabilité sur les réels. Pour chaque valeur réelle, on estime la probabilité que ce type de mitraille ait une teneur plus petite que cette valeur.

2.3 Modélisation du problème inverse

2.3.1 Conventions supplémentaires de notation

Suite aux insuffisances de la solution du problème direct, introduisons des conventions supplémentaires de notation. Celles-ci viennent simplement s'ajouter aux premières. Nous les avons mises à part, car elles sont propres au problème inverse.

$\text{tr}(m,e,c)$ = la teneur (réelle) du type de mitraille m en l'élément e pour la charge c .

$D \equiv$ l'ensemble des répartitions de probabilité sur \mathcal{R} .

$\otimes =$ « une » opération de « multiplication » de signature : $\mathcal{R}^{C \times M} \times D^M \rightarrow \mathcal{R}^M$

2.3.2 Hypothèses de production

Comme il vient d'être dit, la solution proposée au problème direct n'est pas satisfaisante car les différents types de mitraille ne sont pas homogènes en terme de teneur. Nous désirons retrouver « les intervalles » de variation de la teneur de chaque type de mitraille en chaque élément sur l'ensemble des charges. Bien que ce problème soit lié au problème direct, il n'en est pas moins indépendant. Les hypothèses que posées ici jouent sur les mêmes paramètres mais sont différentes que celle du problème direct. Il faut oublier celles posées pour le problème direct et tenir compte de celles-ci.

Les C charges ont été produites

- $\forall m \in \mathcal{M} \quad \forall c \in \mathcal{C}, q(c,m) \in \mathcal{R} (\Rightarrow Q \in \mathcal{R}^{C \times M})$ et est connue. C'est la quantité du type de mitraille m réellement utilisée lors de la production. Elle a été déterminée comme solution du problème direct, bien que l'on en connaisse les insuffisances.
- $\forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}, a(e,c) \in \mathcal{R} (\Rightarrow A_e \in \mathcal{R}^C)$ et est connue. C'est l'analyse mesurée en sortie du circuit de production. Elle est souvent différente de l'analyse désirée utilisée pour le problème direct.

Nous pouvons représenter la teneur de chaque type de mitraille de deux manière :

1. On considère la teneur de chaque type de mitraille en chaque élément pour chaque charge. Dans ce cas, on admet que cette valeur puisse être réelle.
 $\forall m \in \mathcal{M} \forall e \in \mathcal{E} \forall c \in \mathcal{C}, \text{tr}(m,e,c) \in \mathbb{R}$ et est inconnue
2. On considère la teneur de chaque type de mitraille en l'élément e . Cette valeur n'a pas de sens en tant que réel. Nous utilisons donc un concept probabiliste.
 $\forall m \in \mathcal{M} \forall e \in \mathcal{E}, t(m,e) \in \mathbf{D} (\Rightarrow T_e \in D^{\mathcal{M}})$ et est inconnue

2.3.3 Enoncé formel dans le cas 1

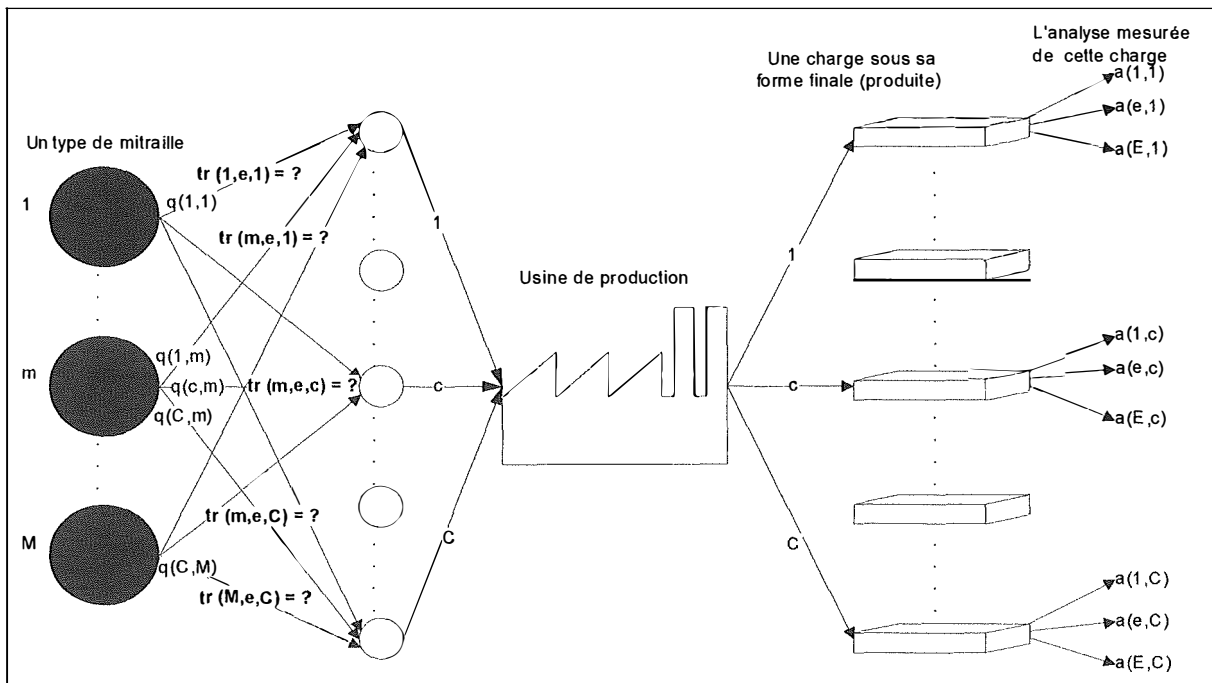
Nous donnons ici l'énoncé du problème inverse, dans le cas où on utilise la première interprétation des teneurs.

Déterminer les $\text{tr}(m,e,c)$ sachant que

$$\sum_{m \in \mathcal{M}} \text{tr}(m,e,c) \times q(m,c) = a(e,c) \quad \forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}$$

« Déterminer la teneur de chaque type de mitraille en chaque élément pour chaque charge produite, connaissant la quantité de tous les types de mitraille utilisées sur l'ensemble des charges, la composition exacte de chaque charge en chaque élément, et la relation qui les lie »

Voici une représentation graphique de ce problème.



Ce problème admet une infinité de solutions mathématiques, car il s'agit d'un système de $E \times C$ équations à $M \times E \times C$ inconnues. Nous sommes certains que la solution physico-chimique qui nous intéresse, c.-à-d. la teneur de chaque type de mitraille en chaque élément pour chaque charge, se trouve parmi ces solutions mathématiques. Malheureusement, nous

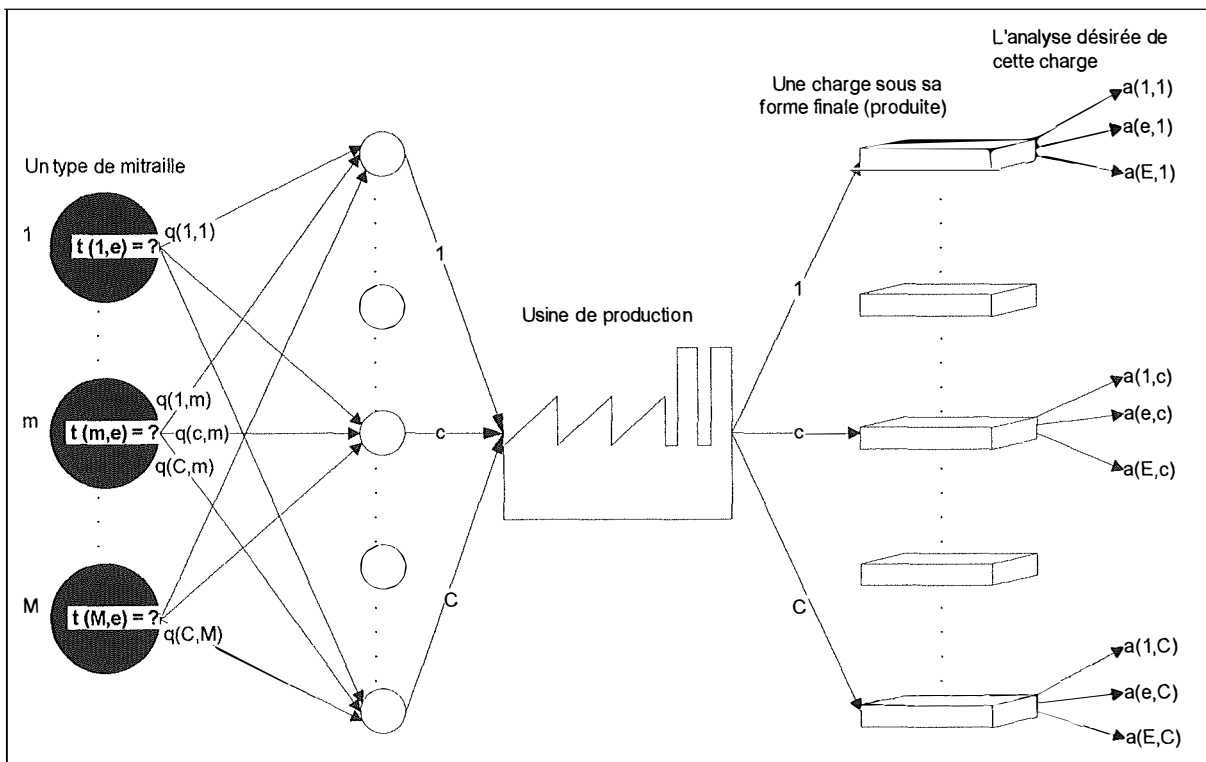
ne pouvons pas la mettre en évidence. En utilisant la première interprétation des teneurs, nous ne sommes donc pas capables de résoudre le problème inverse. Il nous faut donc être moins exigeant quant à la précision d'une solution au problème inverse. Analytiquement, il nous faut diminuer le nombre de variables.

2.3.4 Enoncé formel dans le cas 2

Il est donc apparemment impossible d'utiliser la notion de teneur d'un type de mitraille en un élément pour une charge. Il faut manipuler une autre conceptualisation d'une teneur. Nous donnons ici l'énoncé du problème inverse, dans le cas où on utilise la deuxième interprétation des teneurs.

$$\forall e \in \mathcal{E}, \text{déterminer } T_e \text{ TQ } Q \otimes T_e = A_e$$

Voici une représentation graphique de ce problème.



Dans ce modèle, le fait qui n'existe pratiquement jamais de solution vérifiant l'égalité $Q \times = A_e$ (dans \mathfrak{R}^C , premier cas), est exprimé par l'utilisation de répartitions de probabilité sur les réels. Dire que la teneur du type de mitraille m en l'élément e est représentée par $[T_e]_m = t(m,e)$ signifie que pour tout couple (t_{min}, t_{max}) de valeurs réelles, on connaît la probabilité que la teneur du type de mitraille m en l'élément e sur l'ensemble des charges soit comprise entre t_{min} et t_{max} . Il est évident que l'opérateur de multiplication \otimes n'est pas la multiplication de matrices à composantes réelles, mais en est une extension « adéquate ». Nous supposons qu'il existe ...

2.3.5 Solution actuelle dans le cas 2

Ce problème n'est pas résolu de manière générale. On ne retrouve pas pour chaque type de mitraille m et pour chaque élément e , la répartition de probabilités sur les réels modélisant la teneur en l'élément e du type de mitraille m . On se contente de trouver un seul intervalle $IT(m,e)$ (un seul couple $(tmin,tmax)$), dit de confiance à 95%, pour la teneur de cette mitraille m en l'élément e . On espère ainsi que chaque intervalle $IT(m,e)$ sera représentatif de la variation de la teneur du type de mitraille m en l'élément e sur l'ensemble des charges. Autrement dit, on espère que la teneur en l'élément e du type de mitraille m soit dans cette intervalle avec une probabilité de 0.95. La difficulté de l'énoncé général ci-dessus est ainsi « coupée ». On pourrait utiliser une autre coupe, par exemple en demandant un degré de confiance plus grand.

La société procède en deux parties.

- $\forall e \in \mathcal{E}$, déterminer $T_e \in \mathbb{R}^M$ le vecteur dit des teneurs

$$SC \quad \| Q T_e - A_e \|_2 = \text{Min} \{ \| Q x - A_e \|_2 \mid x \in \mathbb{R}^M \}$$

$\forall m \in \mathcal{M}$, $[T_e]_m$ est le centre de $IT(m,e)$. Q est appelée la matrice des variables d'entrée, A_e le vecteur des variables de sortie, et T_e le vecteur des variables indépendantes.

- $\forall e \in \mathcal{E}$, $\forall m \in \mathcal{M}$ déterminer le rayon de $IT(m,e)$

Le résultat de ces deux parties, sont fournis en sortie d'une routine FORTRAN, qui calcule une estimation de T_e et une erreur standard sur les M composantes du vecteur T_e . Nous ne mettons nullement en doute l'exactitude des calculs, mais à nouveau il semble que cette solution ne convienne pas pour la résolution du problème inverse.

2.3.6 Insuffisances de la solution actuelle

La première objection que l'on peut formuler est que les solutions obtenues dépendent du choix de la norme vectorielle. L'utilisation d'autres normes que $\| \cdot \|_2$ est possible mais souvent plus compliquée.

Outre cette constatation, cette solution est implémentée sous forme de prototype. Elle n'a cependant jamais été utilisée en production, et ce pour deux raisons majeures.

- Les variables d'entrée (Q), soi-disant connues précisément, ne le sont en fait qu'approximativement. En effet, les quantités à utiliser pour la production sont déterminées par la solution du problème direct. Si celle-ci révèle à l'ouvrier de production qu'il doit utiliser 25 tonnes du type de mitraille 27 pour la charge 5032, celui-ci puise dans la cuve du type de mitraille 27, 25 tonnes de mitraille. Or, la façon dont les types de mitraille sont pesées, implique nécessairement que le poids mesuré de la mitraille puisée soit souvent différent de son poids réel. En effet, pour le déterminer, il pèse d'abord la cuve dans laquelle la mitraille va être déposée. Ensuite, il dépose à l'aide d'une grue la mitraille dans la cuve et enfin calcule la différence des deux poids mesurés. Il arrive parfois que l'on pèse la cuve contenant

la mitraille avant que la grue n'ait lâché complètement la mitraille, ou encore lorsque la grue repose encore en partie dans la cuve. Le poids mesuré est donc faussé. Tantôt il est trop grand, tantôt il est trop petit ! Ainsi, l'ouvrier puise peut-être réellement 23 ou 26 tonnes, alors qu'il croit en puiser 25 tonnes. Une multitude d'autres aléas de la production industrielle nous obligent à considérer que le poids mesuré et le poids réellement utilisé diffèrent. Autrement dit, le poids utilisé de chaque type de mitraille n'est connu qu'approximativement. Nous supposons une erreur maximale de 5% ! Les hypothèses du modèle, selon son énoncé probabiliste, ne sont donc pas vérifiées. Il y a donc peu de chance qu'il nous donne les bons résultats.

- Même si les variables d'entrée (Q) étaient exactement connues, ce qui ne sera jamais le cas, le comportement connu d'un moindre carré nous pousse à croire que cette solution ne conviendrait pas. En effet, elle nous renverrait des intervalles de confiance non représentatifs des variations des teneurs de chaque type de mitraille en chaque élément. Par exemple, supposons que parmi tous les types de mitraille, seule une ait une teneur variable et toutes les autres une teneur constante. Si les intervalles de confiance, correspondant à la teneur de chaque type de mitraille en un élément e, étaient représentatifs du comportement réel des teneurs de chaque type de mitraille en cet élément e, on s'attendrait à une solution du type suivant. Les intervalles correspondant aux types de mitraille non variables, seraient de longueur nulle et centrés en la teneur qui leur correspond. L'intervalle correspondant au seul type de mitraille variable, correspondrait au domaine de variation de sa teneur (à 95% près). Nous pensons que la solution trouvée répartirait plutôt l'erreur sur tous les types de mitraille, et donc renverrait un ensemble d'intervalles de confiance, de longueur non nulle, tous non représentatifs du comportement de la teneur de chaque type de mitraille. Ceci est d'ailleurs un comportement classique très connu pour un moindre carré. La solution trouvée ne vérifierait donc pas, selon nous, les spécifications du problème. Nous utilisons le conditionnel, car nous ne sommes pas encore en mesure de le vérifier, étant donné que les variables d'entrée sont incorrectes.

2.4 Le problème de planification de production

Comme nous l'avons déjà expliqué, ce problème consiste à gérer le parallélisme de production. L'usine produit plusieurs charges en même temps. Celles-ci se trouvent toutes à des étapes différentes dans le circuit de production. La société profilARBED dispose déjà d'une application informatique gérant, en temps réel, ce parallélisme. Cette solution ne comporte aucune insuffisance majeure, que ce soit d'un point de vue informatique (algorithmique, interface, gestion de base de données, ...), ou du point de vue des résultats industriels.

Chapitre 3

Introduction à la théorie des ensembles flous

3. Introduction à la théorie des ensembles flous

3.1 Généralités

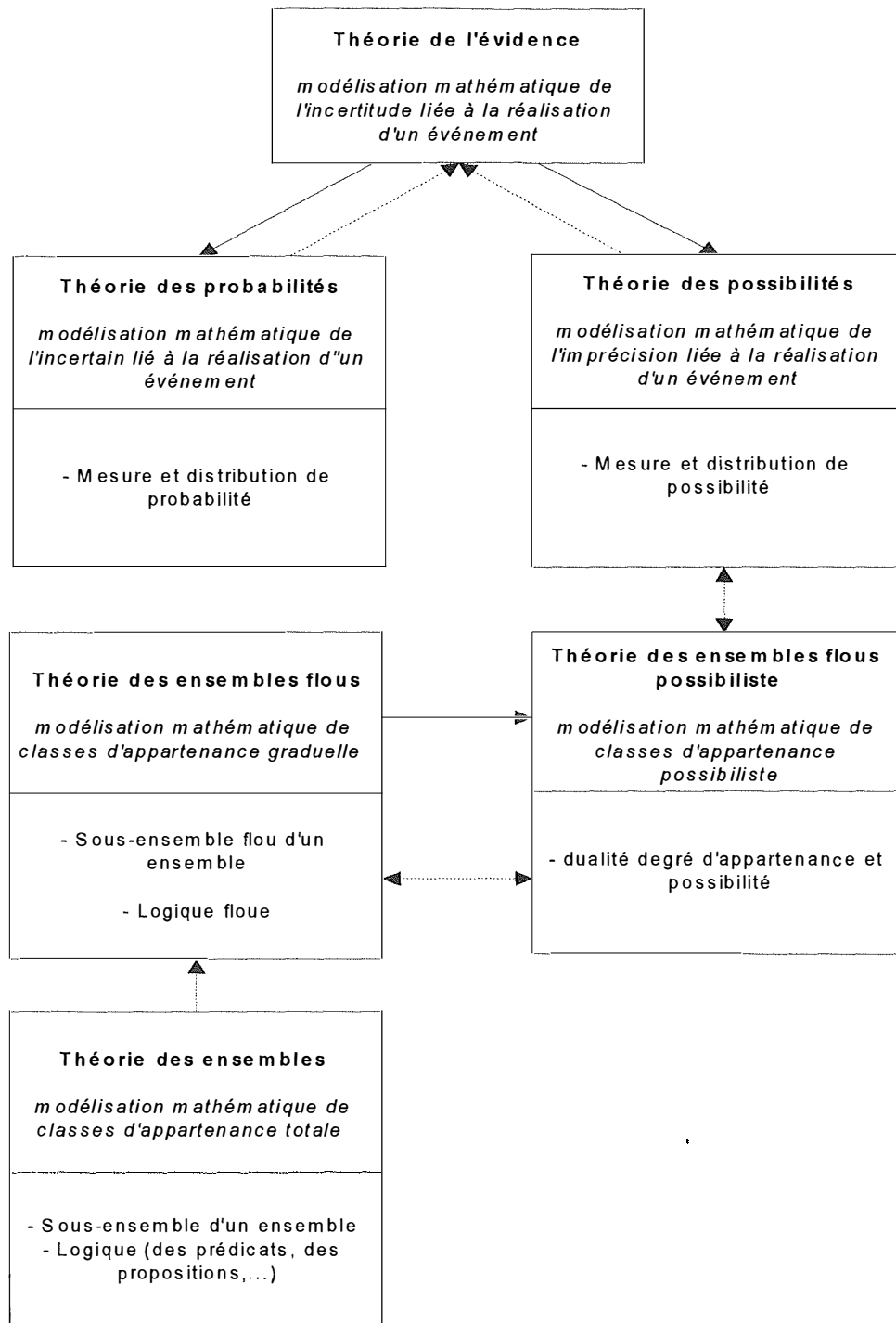
Lorsque l'on s'intéresse à la théorie des ensembles flous, et que l'on commence à lire à ce sujet, on remarque d'emblée que de nombreux concepts et théories mathématiques sont référencés. C'est ainsi que les termes « théorie des probabilités », « théorie des possibilités », « théorie des ensembles flous », « mesure de probabilité », « de possibilité », « logique floue » et des dizaines d'autres sont mélangés. Le point commun entre toutes ces notions, est qu'elles prétendent traiter de « l'incertain » et de « l'imprécis ». Il est parfois difficile de faire le tri et de comprendre la différence fondamentale entre chacun d'eux. De plus, ces différents concepts sont souvent « expliqués » uniquement à travers des définitions mathématiques, qui comme chacun le sait, n'ont aucune signification concrète. Elles ne sont en fait qu'une abstraction formelle, parmi d'autres possibles, de concepts intellectuellement admis par l'esprit et le langage humain, souvent fort ambigu et différent selon les personnes.

Par exemple il est possible « d'expliquer » à quelqu'un ce qu'est une probabilité, en lui fournissant en bloc la formulation mathématique des différents axiomes lui étant associés, sans lui expliquer leurs origines fondamentales. Il faut savoir que, historiquement, cette définition a été fortement controversée avant de s'imposer. Le concept intellectuel qui lui est sous-jacent est basé sur le désir de modéliser la « chance » qu'un événement puisse se réaliser. Or, la notion de « chance de se réaliser » est difficile à percevoir. Elle peut varier selon les personnes. Celle actuellement admise de manière générale, est basée sur une interprétation dite fréquentiste de l'occurrence d'un événement. Ceci est expliqué plus en détail dans la partie sur la théorie des probabilités.

La réaction du profane face à la théorie des ensembles flous, est souvent de ne pas trop y croire et de plutôt la considérer comme une mode, comme un artifice. Dès lors, le but de ce chapitre est de plusieurs ordres. Nous distinguons trois grandes théories qui prétendent traiter de l'incertitude : la théorie des probabilités, la théorie des possibilités, la théorie des ensembles flous. Nous désirons les situer les unes par rapport aux autres, de manière claire et simple, en essayant de percevoir les origines de pensée dont elles dérivent. Nous désirons aussi montrer qu'elles s'appuient toutes sur des bases théoriques formelles et interprétables (du moins autant que la notion de probabilité), contrairement à ce que l'on pourrait croire. En particulier pour la théorie des ensembles flous, nous introduisons les différentes définitions et propriétés dont nous aurons besoin pour comprendre la suite du mémoire. Pour les personnes qui ont déjà des notions claires de cette théorie et qui la situent bien par rapport aux autres, ce chapitre peut évidemment être passé.

Afin de réaliser ces objectifs et avant de rentrer dans les détails, nous pensons qu'il est intéressant de donner une idée globale des liens entre chacune des théories énoncées ci-dessus. Ceci est l'objectif recherché par le schéma ci-dessous. Nous ferons juste après quelques commentaires à son propos. Chaque théorie est représentée par un rectangle. Dans la partie supérieure d'un rectangle se trouve le nom de la théorie. En dessous du nom de la théorie, se trouve une brève description du concept intellectuel qui lui est sous-jacent, c.-à-d. dont elle est « historiquement » issu. Dans la partie inférieure d'un rectangle sont indiquées les principales notions introduites dans la théorie lui correspondant. Une flèche continue d'un rectangle A vers un rectangle B signifie que la théorie représentée par le rectangle A est

une généralisation de la théorie de représentée par le rectangle B. Autrement dit, la théorie représentée par le rectangle B est une particularisation de la théorie représentée par le rectangle A. Une flèche discontinue d'un rectangle A vers un rectangle B signifie que, historiquement, la théorie représentée par le rectangle A est apparue avant la théorie représentée par le rectangle B. Une double flèche discontinue entre deux rectangles signifie que les théories respectives sont apparues plus ou moins en même temps.



3.1.1.1 Quelques premiers commentaires

Remarquons d'emblée qu'une théorie généralisant une autre théorie lui est souvent historiquement postérieure. Bien que ce schéma ne permette pas encore de comprendre fondamentalement chacune des théories, nous espérons qu'il permettra déjà de structurer notre pré connaissance, afin d'orienter plus finement notre compréhension dans les parties qui suivent.

Tout d'abord, nous pouvons déjà dire que le concept intellectuel d'incertitude, de la manière dont il est abordé par la théorie des probabilités, est parfois inadéquat comme nous le montrerons plus loin. C'est la raison pour laquelle il a été reconsidéré et abordé d'une autre manière par Zadeh. La phrase la plus rencontrée dans la littérature dit que « la théorie des probabilités traite de l'incertain, tandis que la théorie des possibilités traite de l'imprécis ». Nous espérons pouvoir vous faire sentir la différence. Etant donné que ces deux théories ont des racines communes, les mathématiciens se sont évertués à les unifier par la théorie de l'évidence. Son mérite est de confirmer le parallélisme entre les deux théories mais n'a, à notre connaissance, aucun intérêt pratique. Dans un premier temps, nous présenterons donc chacune des deux théories indépendamment.

Nous constatons aussi sur le schéma que la théorie des possibilités et la théorie des ensembles flous sont apparemment indépendantes. En fait elles ont été développées en même temps dans le cadre de la théorie possibiliste des ensembles flous. Le concept intellectuel de degré d'appartenance, sur lequel se base la théorie des ensembles flous, est difficilement perceptible. Cependant, il l'est moins si l'on le mixe avec le concept de possibilité sur lequel se base la théorie des possibilités. C'est la raison principale pour laquelle on ne parle pas en général, dans la littérature, de la théorie des ensembles flous, sans parler de la théorie des possibilités, et vice-versa. La façon dont les opérateurs sur les ensembles classiques sont généralisés sur les ensembles flous, est d'ailleurs modelable selon ce que l'on désire lui faire dire, mais prend un sens possibiliste si l'on se place dans le contexte de la théorie des possibilités.

3.2 La théorie des probabilités et la théorie des possibilités

3.2.1 Un mot sur l'algèbre des événements

Soit X l'ensemble de tous les résultats élémentaires possibles d'une expérience. Supposons que X soit fini, ce qui n'est pas toujours le cas.

Ex1 : $X \equiv \{1,2,3,4,5,6\}$ pour l'expérience consistant à jeter un dé une seule fois

Ex2 : $X \equiv \{(1,1), \dots, (5,6), (6,6)\}$ pour l'expérience consistant à jeter un dé deux fois de suite.

Ex3 : $X \equiv \{(AsCoeur, RoiCoeur), (AsCoeur, AsPique), \dots\}$ pour l'expérience consistant à tirer une carte dans un jeu de 52 cartes, puis en tirer une seconde dans le même jeu sans remettre la première.

Un événement élémentaire de cette expérience est modélisé par un singleton contenant un élément de X .

Un événement lié à cette expérience, dans un sens plus général, est une propriété d'une réalisation de cette expérience. Il est modélisé par un sous-ensemble de X . Il est toujours une composition d'opérations « logiques » sur les événements élémentaires. Les opérations de base sont le « ou » (non exclusif) entre deux événements, le « et » entre deux événements, le contraire d'un événement.

L'événement dit certain est modélisé par X .

Ex1 : La propriété de l'événement $X \equiv \{1,2,3,4,5,6\}$ peut être vue comme « on a obtenu 1 ou 2 ou 3...ou 6 lors du jet d'un dé ». Cette événement « est dit certain car on est certain » qu'il va se réaliser (à moins que le dé ne reste en équilibre sur un coin).

L'événement dit impossible est modélisé par \emptyset .

Ex1 : La propriété de l'événement \emptyset peut être vue comme « on a obtenu ni 1, ni 2, ni 3..., ni 6 lors du jet d'un dé ». Cet événement « est dit impossible car on est certain » qu'il ne se réalisera jamais.

L'événement résultat du « ou » entre deux événements, signifie qu'il a la propriété d'au moins un des deux événements de départ. Il est modélisé par un sous-ensemble de X obtenu par l'union des sous-ensembles modélisant les événements de départ.

Ex1 : l'événement qui a la propriété suivante ; « on a obtenu un 1 » ou « on a obtenu un 2 » est modélisé par le sous-ensemble de $X \equiv \{1\} \cup \{2\} \equiv \{1,2\}$ qui est bien l'événement qui a la propriété « on a obtenu 1 ou 2 ».

Ex2 : l'événement qui a la propriété suivante ; « on a obtenu un 1 lors du premier jet et un 6 lors du second jet » ou « on a obtenu un 2 lors du premier jet et un 6 lors du second jet » est modélisé par le sous-ensemble de $X \equiv \{(1,6)\} \cup \{(2,6)\} \equiv \{(1,6), (2,6)\}$, qui est bien l'événement qui a la propriété « on a obtenu 1 ou 2 au premier jet et 6 au second jet » équivalente .

L'événement résultat du « et » entre deux événements, signifie qu'il a la propriété des deux événement de départ. Il est modélisé par un sous-ensemble de X obtenu par l'intersection des sous-ensembles modélisant les événements de départ.

Ex1 : L'événement qui a la propriété suivante ; « on a obtenu un 1 » et « on a obtenu un 2 », un événement évidemment impossible, est modélisé par le sous-ensemble de $X \equiv \{1\} \cap \{2\} \equiv \emptyset$ qui est bien la modélisation de l'événement impossible.

L'événement contraire d'un autre événement, signifie qu'il a la propriété contraire de l'événement de départ. Il est modélisé par un sous-ensemble de X obtenu par complémentarité du sous-ensemble modélisant l'événements de départ.

Ex1 : L'événement qui a la propriété contraire à la propriété « on a obtenu un 1 », est modélisé par le sous-ensemble de $X \equiv \{1\}^C \equiv \{2,3,4,5,6\}$, qui est bien la modélisation de l'événement ayant la propriété «on n'a pas obtenu 1 ».

Deux événements sont dits disjoints ssi l'intersection de leurs modèles ensemblistes est vide.

Deux événements sont dits non interactif ssi la réalisation de l'un « n'influe pas » sur la réalisation simultanée des deux ».

Ex3 : l'événement A qui a la propriété « obtenir un cœur au second tirage » et l'événement B qui a la propriété suivante « obtenir un cœur au premier tirage » sont interactifs lorsqu'il n'y a pas de remise. En effet, l'obtention d'un cœur au second tirage est influencée par l'obtention ou non d'un cœur au premier tirage.

Ce concept est défini par la notion d'indépendance en théorie des probabilités. Il garde le même nom en théorie des possibilités.

Soit $P(X)$ l'ensemble des parties de X . $P(X)$ n'est autre que la modélisation de l'ensemble de tous les événements (pas nécessairement élémentaire) possibles liés à l'expérience dont les événements élémentaires sont modélisés par X .

Ex1 : $P(X) \equiv \{\emptyset, \{1\}, \dots, \{6\}, \{1,2\}, \dots, \{5,6\}, \{1,2,3\}, \dots, \{4,5,6\}, \dots, \{2,3,4,5,6\}, X\}$

Par abus de langage, nous ne ferons parfois plus la différence entre un événement et son modèle ensembliste.

3.2.2 Les fondements de la théorie des probabilités

3.2.2.1 Le concept intellectuel sous-jacent

Cette théorie a pour but de modéliser l'incertitude sur la réalisation d'un événement, par ce que l'on appelle l'interprétation fréquentiste d'une probabilité. Celle-ci se base sur certains choix.

La probabilité d'un événement lié à une expérience est « l'idéalisation » de sa fréquence de réalisation sur un grand nombre de réalisations de cette expérience. Ainsi, la probabilité d'obtenir « face » lors du jet d'une pièce, est une idéalisation du nombre de fois que l'on obtiendrait « face », divisé par le nombre de jets, si l'on réalisait un très grand nombre de jets de cette pièce (mathématiquement infini). K. Pearson a jeté 24000 fois une pièce, et a obtenu 12012 fois « face ». Ainsi l'on convient que la probabilité d'obtenir face lors du jet d'une pièce vaut $\frac{1}{2}$. Remarquons qu'elle aurait très bien pu valoir 0.9 si la pièce était « truquée ». La notion de mesure de probabilité que nous allons introduire et qui est enseignée dans toutes les écoles, n'est autre qu'une modélisation mathématique la plus cohérente possible avec cette interprétation.

Si on arrive à formaliser la notion de probabilité de manière cohérente avec cette interprétation, alors on peut se servir des probabilités pour modéliser certains systèmes de manière à ce que les résultats et implications trouvés à partir du modèle en terme de probabilité, puissent être interprétés selon cette interprétation fréquentiste. Par exemple, supposons que l'on modélise la teneur d'une charge en fer en terme de probabilité. Si le modèle nous permet de conclure « la probabilité que la teneur en fer de cette charge soit inférieure à 85% vaut 0.99 », on espère que cela signifie que si l'on avait produit 10000 charges dans les mêmes conditions, on en aurait obtenu plus ou moins 9990 avec une teneur en fer inférieure à 85%. On en conclurait donc qu'il y aurait alors de « fortes chances » pour que la teneur en fer de cette charge soit inférieure à 85%. En prenant un « faible risque », on irait même jusqu'à dire que la teneur en fer de cette charge est inférieure à 85%...

3.2.2.2 Les limites pratiques de la théorie des probabilités

On sent nettement dans cet exemple en quoi l'interprétation fréquentiste d'une probabilité peut être, à raison, controversée. De plus tous les événements présents dans la nature ne se prêtent pas toujours à une interprétation fréquentiste (probabiliste). En effet, supposons que l'on veuille calculer la probabilité que le patron du groupe international ARBED soit présent aux FUNDP le mercredi 4 septembre. Un raisonnement probabiliste nous conduirait à compter le nombre de mercredi que ce monsieur a passé aux FUNDP pendant sa vie. Ou, peut-être, devrions nous nous limiter aux années de sa présidence ? Ou encore sur tous les premiers mercredi du mois ? On voit là, que ce genre d'événement ne se prête pas du tout à un raisonnement probabiliste. Par contre en ce qui concerne le jet d'une pièce, nous pouvons raisonnablement penser qu'il s'y prête bien. C'est ainsi que nous sommes naturellement amenés à différencier les événements probabilistes et les événements non probabilistes. La théorie des probabilités peut donc servir à modéliser de manière satisfaisante des systèmes régis par des événements de nature probabiliste. L'expérience nous montre qu'il y en a beaucoup, ne fut-ce que par l'utilisation importante de cette théorie dans le monde.

3.2.2.3 Les axiomes et définitions de base dans le cas discret

Comme nous avons déjà essayé de le faire sentir, une probabilité a pour but d'associer à chacun des événements possibles d'une expérience, un nombre réel compris entre 0 et 1, cohérent avec l'interprétation fréquentiste d'une probabilité. Ainsi, pour rester cohérent avec cette interprétation, on désire que

- la probabilité de l'événement certain soit égale à 1.
- la probabilité de l'événement impossible soit égale à 0.
- la probabilité de l'événement A ou B vaille la somme des probabilités des événements A et B pris seuls lorsqu'ils sont disjoints.
- la probabilité de l'événement A et B vaille le produit des probabilités de A et B pris seuls lorsque A et B sont non interactifs.

La notion de non-interactivité d'événement est formalisée par la notion d'indépendance.

Une mesure de probabilité **P** sur **X** doit donc, par définition, respecter les axiomes suivants :

$P : P(X) \longrightarrow [0,1] \subseteq \mathfrak{R}$ <p>vérifiant les axiomes</p> $P(X) = 1$ $P(\emptyset) = 0$ $\forall A, B \in P(X), A \cap B \equiv \emptyset \Leftrightarrow P(A \cup B) = P(A) + P(B)$ $\forall A, B \in P(X), A \text{ indépendant de } B \Leftrightarrow P(A \cap B) = P(A) \times P(B)$

Toute application vérifiant ces axiomes est appelée mesure de probabilité sur X.

3.2.2.4 Dualité mesure de probabilité-distribution de probabilité

Comme nous le devinons facilement, la probabilité d'un «événement» d'une expérience modélisée par X peut toujours s'exprimer comme une composition d'opérations arithmétiques sur les mesures de probabilités des événements élémentaires de cette expérience. Finalement, seule la probabilité des événements élémentaires est importante. C'est la raison pour laquelle toute fonction de la forme

$$p : X \longrightarrow [0,1] \subseteq \mathfrak{R}$$

vérifiant la propriété

$$\sum_{x \in X} p(x) = 1$$

est appelée **distribution de probabilité sur X**.

Une distribution de probabilité p définit une mesure de probabilité et vice-versa. Connaître la mesure de probabilité d'une expérience ou sa distribution revient à connaître toute l'information probabiliste sur l'expérience.

3.2.2.5 L'évaluation d'une mesure de probabilité

On ne connaît pas toujours la probabilité de tous les événements élémentaires. Pratiquement, cela ne nous empêche pas nécessairement d'évaluer la probabilité d'un événement (non élémentaire) de cet expérience. Cela est dû au fait que les probabilités d'événements différents sont liées arithmétiquement.

Ex : Soit X l'ensemble des événements élémentaires d'une expérience consistant à jeter un dé pipé. Si nous savons que $P\{1\} = 0.6$ et que $P\{1,6\} = 0.7$ alors nous pouvons en déduire que $P\{6\} = 0.1$ et que $P\{2,3,4,5\} = 0.3$.

Si nous connaissons la probabilité d'un événement, on connaît automatiquement celle de son complémentaire, car ces deux probabilités sont liées par une relation arithmétique. L'incertitude liée à un événement probabiliste est entièrement décrite par la probabilité de cet événement.

3.2.2.6 *La suite de la théorie des probabilités*

Maintenant que les bases de la théorie des probabilités ont été introduites, nous voulons dire un dernier mot sur le reste du développement de cette théorie.

Nous avons considéré jusqu'à présent que X est un ensemble fini. Ce n'est pas toujours le cas comme nous l'avons annoncé. X peut être un ensemble contenant une infinité d'éléments. Par exemple, pour définir une mesure de probabilité à propos de la position d'arrivée d'une flèche dans une cible, X est alors modélisé par une surface finie de \mathbb{R}^2 . Un événement élémentaire est alors modélisé par un couple (x,y) de réels. La probabilité qu'une flèche tombe en (x,y) serait nulle par cohérence avec l'interprétation fréquentiste. On remarque donc que le passage au cas où X n'est pas fini n'est pas évident. C'est la raison pour laquelle on introduit non plus la notion de distribution de probabilité, mais plutôt la notion de **répartition de probabilité**, une sorte de distribution des « probabilités cumulées ».

On introduit aussi les notions de **variables aléatoires** (continues ou discontinues), pour exprimer certains comportements probabilistes, de manière indépendante du contenu de X . Certains systèmes, bien qu'ayant un ensemble différent d'événements, peuvent avoir le même comportement probabiliste. Par exemple, on peut imaginer le passage des voitures à un péage sur les autoroutes, est « similaire » en terme de probabilité sur le tronçon Paris-Poitiers que sur le tronçon Barcelone-Taragone. En terme de probabilité, on passe à un niveau d'abstraction supérieur en modélisant le comportement probabiliste à un péage, sur un tronçon d'autoroute, en général. On peut encore monter d'un cran dans l'abstraction, en imaginant que le comportement probabiliste du passage des voitures sur un tronçon d'autoroute, est « similaire » à celui du passage de personnes aux barrières de l'entrée d'une rame de métro. On peut ainsi monter en abstraction pour en arriver finalement à la description d'une variable aléatoire de Poisson. Il s'agit d'une distribution de probabilité sur les réels, permettant de modéliser le comportement probabiliste de chacun des systèmes pris en exemple, et probablement encore d'autres. L'idée générale est de considérer une variable aléatoire comme la projection de X sur l'ensemble des réels. On parle alors de distribution de probabilité d'une variable aléatoire. Il existe ainsi plusieurs variables aléatoires types, chacune étant l'image d'un comportement probabiliste particulier. Ainsi, la variable aléatoire uniforme sert à modéliser tout système se comportant de telle manière que les événements élémentaires soient équiprobables (ce qui n'est évidemment pas toujours le cas). Par exemple, dans les sciences physiques, les variables aléatoires de Dirac, de Gauss, de Poisson et beaucoup d'autres sont fortement utilisées.

Nous nous arrêterons ici quant à notre discussion sur les fondements de la théorie des probabilités, car l'étude des probabilités est très vaste, et là n'est pas notre objectif.

3.2.3 Les fondements de la théorie des possibilités

3.2.3.1 Pourquoi la théorie des possibilités ?

Nous avons vu que la théorie des probabilités est une manière de modéliser la notion d'incertitude sur la réalisation d'un événement, à travers une interprétation fréquentiste de sa réalisation. Nous avons aussi remarqué que cette interprétation, et donc cette théorie, n'est pas toujours adéquate selon ce que l'on veut modéliser. Elle ne convient que pour modéliser des systèmes où les événements sont dits « probabilistes » ou « aléatoires » au sens étymologique du terme. Tous les événements ne le sont pas, comme par exemple celui du directeur. On les appelle **événements « non probabilistes »**, simplement parce que l'on ne dispose pas de moyen d'évaluer leur probabilité à partir de l'interprétation fréquentiste.

Cette première approche est donc intéressante mais insuffisante. Zadeh l'a bien compris, et c'est pour cette raison qu'il tente une deuxième approche liée à la notion d'événement « possible ». Sa « théorie des possibilités » est introduite en 1978, en complément de sa « théorie des ensembles flous » introduite en 1965, que nous aborderons dans la partie suivante.

3.2.3.2 Le concept intellectuel sous-jacent

Cette théorie a pour but de modéliser l'incertitude sur la réalisation d'un événement non probabiliste, par une interprétation très subjective de possibilité d'un événement. Ainsi l'on attribue à chacun des événements d'une expérience à résultats non probabilistes, une possibilité comprise entre 0 et 1, selon notre interprétation des informations que l'on dispose à leur sujet.

Pour reprenons notre exemple du président de l'ARBED énoncé dans la partie sur les limites pratiques de la théorie des probabilités. Si l'on se place du point de vue de sa secrétaire, elle nous dira qu'il est impossible qu'il soit présent aux FUNDP de Namur le mercredi 4 septembre, parce qu'elle sait qu'il a une réunion importante ce jour-là avec notre premier ministre Monsieur Dehaene. Par contre un illustre inconnu interrogé dans la rue, ne connaissant pas du tout le personnage, répondra que ce n'est pas impossible, car après tout ce directeur y a peut-être fait ses études. Pour donner un autre exemple du même acabit. Supposons que l'on soit dans la semaine du 8 au 12 juillet 1996. Essayons d'évaluer la possibilité que, pendant cette semaine, F.P. rencontre M.L. sur le site de Belval « lundi ou mardi ». Un illustre inconnu répondrait que c'est quasi impossible. J.-P. L., lui, répondrait que c'est fort possible parce qu'il sait que F.P. est en stage à profilARBED au service M.L.. M.L. lui-même dirait que c'est quasi certain parce qu'il sait qu'il a rendez-vous avec F.P. le 8 et le 9 juillet 1996 à cet endroit. Madame Soleil, elle, répondrait que c'est certain car elle sait prédire l'avenir ...

Il y a donc bien, dans tous ces exemples, une graduation dans l'évaluation, certes très subjective, de la possibilité de réalisation d'un événement. Cependant, cette subjectivité est tout de même limitée par le fait que l'on désire que la possibilité de l'événement certain vaille 1, pour exprimer le fait qu'il est le plus possible de tous les événements possibles. Par un raisonnement similaire, on désire que le degré de possibilité de l'événement impossible vaille 0.

De plus, il semble intuitif de dire que, si un événement A est fortement possible et qu'un autre événement B est faiblement possible, alors l'événement A ou B est fortement possible avec la même possibilité que A. C'est comme si la possibilité d'un événement A l'emportait sur tout événement B moins possible que A, dans un événement A ou B. Par exemple, plaçons nous du point de vue de J.-P. L.. Supposons que, pour lui, pendant la semaine du 8 au 12 juillet, il est fortement possible que F.P. rencontre M.L. sur le site de Belval « lundi ou mardi », et faiblement possible le week-end (samedi ou dimanche). A la suite de cela, il décide d'attribuer une possibilité de 1 pour une rencontre « lundi ou mardi », et 0.2 pour une rencontre « le week-end ». Intuitivement, il lui semble fortement possible, avec une possibilité de 1, que pendant la semaine du 8 au 12 juillet, F.P. rencontre Monsieur M.L. « le lundi ou mardi » ou « le week-end ». Supposons aussi que J.-P. L. attribue une possibilité de 0.9 au fait qu'il y ait 25° pendant ce week-end. Il lui semble alors logique de dire que la possibilité de rencontre, sur le site de Belval, entre F.P. et M.L., ce week-end à une température de 25° est de 0.2. Tout se passe comme si la possibilité d'un événement A l'emportait sur tout événement B plus possible que A et non interactif avec A, dans un événement A et B.

3.2.3.3 Les axiomes et définitions de base dans le cas discret

Soit X l'ensemble de tous les événements élémentaires non probabilistes possibles d'une expérience. Supposons que X soit fini, ce qui n'est pas toujours le cas.

Ex : $X \equiv \{\text{lundi, mardi, mercredi, jeudi, vendredi, pas de rencontre}\}$ pour la présence de F.P. sur le site de Belval la semaine du 8 au 12 juillet 1996.

Soit $P(X)$ l'ensemble des parties de X. $P(X)$ n'est autre que l'ensemble de tous les événements possibles liés à l'expérience.

Ex : $P(X) = \{\emptyset, \{\text{lundi}\}, \dots, \{\text{vendredi}\}, \{\text{lundi, mardi}\}, \dots, X\}$

Comme nous avons déjà essayé de le faire sentir, une possibilité a pour but d'associer à chacun des événements possibles d'une expérience, un nombre réel compris entre 0 et 1, cohérent avec l'interprétation subjective d'une possibilité, modélisant un degré de certitude de sa réalisation.

Une mesure de possibilité Π sur X , en fonction de cette interprétation, doit vérifier les axiomes suivant :

$$\begin{array}{c} \Pi : P(X) \longrightarrow [0,1] \subseteq \mathfrak{R} \\ \\ \text{vérifiant les axiomes} \\ \\ \Pi(X) = 1 \\ \Pi(\emptyset) = 0 \\ \forall A, B \in P(X), \Pi(A \cup B) = \text{Sup}\{ \Pi(A), \Pi(B) \} \\ \forall A, B \in P(X), A \text{ et } B \text{ non interactifs, } \Pi(A \cap B) = \text{Inf}\{ \Pi(A), \Pi(B) \} \end{array}$$

Remarquons que dans ce cas, les opérateurs Sup et Inf, sont équivalents aux opérateurs Max et Min car nous ne considérons qu'un nombre fini d'événements.

3.2.3.4 La dualité mesure de possibilité-distribution de possibilité

La possibilité d'un «événement d'une expérience modélisée par X » peut toujours s'exprimer comme le supremum de la mesure de possibilité de tous les événements élémentaires qui le compose (par un « ou » logique). Finalement, seule la possibilité des événements élémentaires est importante. C'est la raison pour laquelle toute fonction de la forme

$$\pi : X \longrightarrow [0,1] \subseteq \mathfrak{R}$$

vérifiant la propriété

$$\text{Sup} \{ \pi(x) \mid x \in X \} = 1$$

est appelée **distribution de possibilité sur X** . Une distribution de possibilité π définit une mesure de possibilité et vice-versa.

3.2.3.5 L'évaluation d'une mesure de possibilité

On ne connaît pas toujours la possibilité de tous les événements élémentaires d'une expérience. Pratiquement, contrairement aux probabilités, cela nous empêche d'évaluer facilement la mesure de possibilité d'un événement (non élémentaire) de cette expérience. Cela est dû au fait que les mesures de possibilités d'événements ne sont pas liées arithmétiquement.

La seule propriété que l'on puisse déduire par propriété de l'opérateur Sup est la suivante :

$$\forall A, B \in X \ A \subseteq B \Rightarrow \pi(A) \leq \pi(B)$$

ce qui ne nous avance pas à grand chose.

Plus particulièrement, si l'on connaît la possibilité d'un événement (modélisé par A), on ne peut pas toujours en déduire la possibilité de son contraire (modélisé par A^c). En effet

$$\forall A \in X, \quad \begin{array}{l} \text{Si } \Pi(A)=1 \text{ alors } \Pi(A^c) \leq 1 \\ \text{Sinon } \Pi(A^c)=1 \end{array}$$

On peut donc avoir $\Pi(A^c)$ et $\Pi(A) = 1$. Si un événement est fortement possible, son complémentaire peut l'être aussi. Si on ne connaît pas (toute) la distribution de possibilité de l'expérience, ceci implique une incertitude totale sur un événement A dont on ne connaît que sa possibilité. L'incertitude liée à un événement A non probabiliste n'est pas toujours entièrement décrite par sa possibilité $\Pi(A)$, si l'on n'a pas d'autre information que cette valeur.

3.2.3.6 Mesure de nécessité

C'est la raison pour laquelle, pratiquement, on « complète » la mesure de possibilité d'un événement par une mesure de nécessité N . Les axiomes sont déduits de manière similaire à ceux d'une mesure de possibilité. Il suffit de considérer la notion subjective de nécessité utilisée par notre esprit. Ainsi, **une mesure de nécessité sur X** , vérifie les axiomes suivants :

$$N : P(X) \longrightarrow [0,1] \subseteq \mathfrak{R}$$

vérifiant les axiomes

$$N(X) = 1$$

$$N(\emptyset) = 0$$

$$\forall A, B \in P(X), N(A \cap B) = \inf\{ N(A), N(B) \} = \min\{ N(A), N(B) \}$$

Les mesures de possibilité P et de nécessité N d'une expériences sont duales. Si l'on connaît entièrement l'une, on connaît automatiquement l'autre. Si on connaît une partie de l'une, on connaît la partie « complémentaire » de l'autre. Cette dualité est exprimable mathématiquement par la propriété facilement déductible des axiomes liés aux deux théories.

$$\forall A \in P(X), N(A) = 1 - \Pi(A^c)$$

3.2.3.7 La suite de la théorie des possibilités

Jusqu'à présent, nous avons supposé que l'ensemble X est fini. Comme nous l'avions annoncé, et de manière analogue à ce qui a été dit dans la partie sur la théorie des probabilités, ce n'est pas toujours le cas. Toutes les notions dont nous venons de parler ont été généralisées aux cas où X est un ensemble contenant une infinité d'éléments Nous ne nous étendrons pas sur ce sujet car d'une part, ce n'est pas le but de ce mémoire, et d'autre part nous disposons déjà de toutes les notions nécessaires pour comprendre la théorie possibiliste des ensembles flous. Remarquons tout de même que la suite du développement de cette théorie est typiquement calquée sur celle des probabilités, tout en laissant plus de liberté sur le choix des valeurs modélisant l'incertitude des événements.

3.2.4 Une comparaison des deux théories

La présentation de la théorie des probabilités est volontairement très similaire à celle de la théorie des possibilités. Nous espérons que, au-delà de l'aspect purement mathématique, cela permet de percevoir leurs différences et leurs atouts respectifs. Nous pouvons maintenant comprendre la comparaison entre les deux théories, souvent utilisée dans la littérature. De façon imagée, on peut résumer cette comparaison en disant que la théorie des possibilités est la théorie des probabilités où on remplace les opérateurs « + » et « × » par les opérateurs « Sup » et « Inf » (équivalent à Max et Min dans le cas fini). Nous exprimons ce parallélisme dans le tableau suivant :

Théorie des probabilités	Théorie des possibilités
distribution p et mesure P	distribution π et mesure Π
$x \in X$ ensemble des événements élémentaires probabilistes (supposé fini)	$x \in X$ ensemble des événements élémentaires non probabilistes (supposé fini)
$\forall A, B \in P(X)$	$\forall A, B \in P(X)$
$P(A) = \sum_{x \in A} p(x)$	$P(A) = \text{Max} \{ p(x) \mid x \in A \}$
$P(A \cup B) = P(A) + P(B)$ si $A \cap B \equiv \emptyset$	$\Pi(A \cup B) = \text{Max} \{ \Pi(A), \Pi(B) \}$
$P(A \cap B) = P(A) P(B)$ si indépendance	$\Pi(A \cap B) = \text{Min} \{ \Pi(A), \Pi(B) \}$
$P(A) + P(A^c) = 1$	$\text{Max} (\Pi(A), \Pi(A^c)) = 1$ $\Pi(A) + \Pi(A^c) \geq 1$ $N(A) + N(A^c) \leq 1$

3.3 La théorie des ensembles flous

3.3.1 Le concept intellectuel sous-jacent

Toute personne ayant quelques connaissances en mathématique, sait que le concept d'ensemble a été inventé pour modéliser le regroupement d'entités similaires. C'est ainsi que l'on peut parler de l'ensemble des animaux ou des végétaux. Ce concept est abordé par la théorie mathématique des ensembles. On y définit ainsi une série d'ensembles de base, tels que \mathbb{N} et \mathbb{R} . On y introduit aussi les notions d'appartenance (totale) d'un élément à un ensemble et de sous-ensemble, indissociables l'une de l'autre. Nous ne rentrerons pas dans les détails de cette théorie qui, encore de nos jours, fait l'objet de paradoxes. On définit aussi une série d'opérations et de relations sur les ensembles. Les opérations de base sont l'union et l'intersection. Une relation fort connue entre ensembles est l'égalité d'ensembles.

Grâce à ces notions, les sciences telles que la physique ou l'informatique et des disciplines telles que l'industrie, peuvent modéliser des systèmes allant des plus simples aux plus complexes. Par exemple, la modélisation des âges (en année) des personnes majeures peut se faire, dans une base de donnée par l'ensemble $X \equiv \{x \in \mathbb{R} : 18 \leq x \leq 150\}$. Tout le monde sera d'accord avec cette modélisation. Cependant, de nos jours, ce concept paraît parfois trop rigide. Supposons que nous voulions modéliser les âges des personnes dites « jeunes ». Une personne A le fera avec le sous-ensemble des réels $X_A \equiv \{x \in \mathbb{R} : 18 \leq x \leq 29\}$, tandis qu'une personne B le fera avec un autre sous-ensemble des réels $X_B \equiv \{x \in \mathbb{R} : 16 \leq$

$x \leq 25$ }. D'autres encore essayeront de les modéliser par des sous-ensembles encore différents. Mais, quelle que soit la complexité avec laquelle il définiront leur sous-ensemble, toutes ces personnes n'arriveront jamais à se mettre d'accord.

Nous en concluons que le concept d'ensemble ne suffit plus pour modéliser la « jeunesse ». Selon Zadeh, ceci s'explique par le fait que ce concept est flou (fuzzy), imprécis. Sa sémantique diffère selon les personnes. C'est à la suite de ce genre de problème qu'il développe la théorie des sous-ensembles flous. Par abus de langage, on la nomme aussi théorie des ensembles flous. Elle trouve une « réponse » au problème évoqué ci-dessus, en faisant un raisonnement très simple. On ne peut pas modéliser le concept de « jeunesse » par la théorie des ensembles car, intrinsèquement cette théorie nous oblige à considérer que l'âge x est, ou n'est pas, considéré comme l'âge d'une personne « jeune ». Or, il n'y aura jamais de réponse unanime à propos de tous les âges. Bien que tout le monde soit sûrement d'accord pour dire que l'on est jeune à 20 ans, il n'y a pas de doute que des disputes se déclareront, quant à savoir si l'on est encore considéré comme « jeune » à l'âge de 30 ans.

Afin d'essayer de mettre tout le monde d'accord, Zadeh propose de considérer qu'une personne de 30 ans « est et n'est pas jeune ». Il modifie le concept d'appartenance totale d'un élément à un ensemble et le remplace par le concept d'appartenance partielle. Il propose de considérer que 30, en tant qu'âge d'un jeune, appartient, par exemple, « à moitié » à l'ensemble des âges des personnes en général. On en peut pas dire que cet âge appartient totalement à l'ensemble des âges des jeunes, mais on ne peut pas dire non plus que cet âge n'y appartient pas. Zadeh attribue à la valeur 30, un degré d'appartenance de $\frac{1}{2}$ à l'ensemble $X \equiv \{x \in \mathbb{R} : 0 \leq x \leq 150\}$. Ce dernier représentant, sans faire l'objet de disputes, les âges possibles (en année) des humains. Il gradue ainsi l'appartenance d'un élément à un ensemble par un réel compris entre 0 et 1. La sémantique du degré 0 est la non appartenance, tandis que le 1 est l'appartenance totale au sens commun du terme. Le résultat de l'attribution d'un degré d'appartenance à tous les éléments d'un ensemble, devient ce que l'on appelle un sous-ensemble flou de cet ensemble. Bien que ce genre de raisonnement ne permette sans doute pas non plus de mettre tout le monde d'accord (car il faudra encore discuter de la valeur du degré d'appartenance à un ensemble), on peut au moins espérer qu'il permettra d'arriver plus vite à un compromis. Zadeh étend donc en quelque sorte le concept de sous-ensemble d'un ensemble, au concept de sous-ensemble flou d'un ensemble. Un sous-ensemble d'un ensemble est un cas particulier de sous-ensemble flou de ce même ensemble, où le degré d'appartenance est limité aux deux valeurs 1 et 0. Par abus de langage, dans la littérature, on dit que Zadeh a étendu la notion d'ensemble à la notion d'ensemble flou, ce qui évidemment est faux puisque, par définition, le concept de sous-ensemble flou n'a pas de sens sans le concept d'ensemble. Nous nous permettrons d'utiliser parfois cet abus.

Maintenant que nous avons « étendu » la notion de sous-ensemble d'un ensemble, à la notion de sous-ensemble flou d'un ensemble, il est intéressant d'étendre aux sous-ensembles flous les concepts de base, liés aux (sous-)ensembles. Par exemple, on peut définir l'union de deux sous-ensembles flous d'un ensemble. Il est évident que cette extension se fait de telle manière que si l'on s'y limite aux sous-ensembles flous qui sont des sous-ensembles, la définition de l'union des deux sous-ensembles flous devient équivalente à la définition de l'union des deux sous-ensembles. D'autres questions apparaissent alors. Peut-on parler de l'union de deux sous-ensembles flous de deux ensembles différents ?

Le but de la théorie des ensembles flous est, d'une part de formaliser le concept de sous-ensemble flou d'un ensemble, et d'autre part de trouver une réponse à la plupart des questions qui viennent naturellement à l'esprit. Depuis une vingtaine d'années, elle connaît

un engouement assez prononcé, surtout au Japon. Dans beaucoup de cas, elle permet de résoudre des problèmes de manière plus satisfaisante. Il s'agit surtout de problèmes où les données et concepts à manipuler sont imprécis ou flous, tels que dans la médecine ou dans l'industrie. Remarquons que l'esprit humain raisonne de manière « floue ». Le langage lui-même est flou. Depuis le début de ce mémoire, nous vous parlons avec des mots qui sont évidemment remplis d'ambiguïté, mais ceci n'empêche certainement pas d'avoir l'impression de comprendre ce que nous disons. C'est la raison pour laquelle la théorie des ensembles flous est aussi fort utilisée en linguistique. Cette théorie fait encore actuellement l'objet de beaucoup de recherche. Des chercheurs tentent notamment d'étendre des techniques mathématiques de plus en plus compliquées selon la théorie de ensembles flous. C'est ce que l'on appelle la « fuzzyfication ». On parle par exemple de programmation linéaire floue, de réseau de Petri flous, ...

Avant de rentrer dans les détails mêmes de la théorie, nous pensons qu'il est utile de signaler un point important. La façon dont est étendue la théorie des (sous-)ensembles à la théorie des sous-ensembles flous est évidemment basée sur des choix, préservant la cohérence avec la théorie des ensembles. Ces choix se font notamment sur base de la signification du degré d'appartenance. Celle signification peut varier selon les personnes ou les besoins. Ainsi l'on peut concevoir la théorie des ensembles flous de telle manière qu'elle convienne mieux pour résoudre certains types de problèmes. Il y a donc plusieurs manières de l'étendre. Par exemple, H.J. Zimmermann présente plusieurs manières de considérer le concept d'intersection de (sous-)ensembles flous. Cependant, comme nous le verrons plus tard, le lien entre le concept de degré d'appartenance et de possibilité se fait facilement (théoriquement il n'est pas nécessaire). Suite à cela, cette extension est souvent présentée comme étant unique, et égale à celle où la sémantique du degré d'appartenance est liée à celle de possibilité. Il s'agit du principe de restriction expliqué par Zadeh. On l'appelle restriction car on restreint le sens du degré d'appartenance à celui de possibilité. Dans ce cas particulier, nous parlerons de la théorie possibiliste des ensembles flous. Nous la présentons dans la dernière partie de ce chapitre. Mais dans un premier temps, présentons les axiomes et définitions de base de la théorie des sous-ensembles flous, sans prendre encore cette direction particulière d'extension.

3.3.2 Les définitions de base

3.3.2.1 Les sous-ensembles flous

Soit **X un ensemble** quelconque. On peut caractériser un **sous-ensemble A de X** de deux manières différentes :

La première est une énumération de chaque élément de X qui appartient au sous-ensemble de X. Cette énumération peut se faire élément par élément ou, lorsque c'est possible, par un prédicat.

Ex : Soit \mathbb{N} l'ensemble des naturels. Soit A « quelque chose » qui représente les âges admissibles pour un être humain. On modélise naturellement A par un sous-ensemble de \mathbb{N} de la façon suivante. $A \equiv \{1, 2, 3, \dots, 150\} \equiv \{x \in \mathbb{N} \mid 0 \leq x \leq 150\}$. On énumère tous les naturels faisant partie du sous-ensemble A.

La deuxième manière est une fonction appelée fonction μ_A dite d'appartenance, définie comme suit :

$$\mu_A : X \longrightarrow \{0,1\} \subseteq \mathbb{N}$$

x	$\mu_A(x) = 1$	si x est considéré comme faisant partie de la classe
	$\mu_A(x) = 0$	sinon

Ex : Soit \mathbb{N} l'ensemble des naturels. Soit A « quelque chose » qui représente les âges admissibles pour un être humain. On modélise A par une fonction μ_A dite d'appartenance de la façon suivante :

$$\mu_A : \mathbb{N} \longrightarrow \{0,1\} \subseteq \mathbb{N}$$

x	$\mu_A(x) = 1$	si $0 \leq x \leq 150$
	$\mu_A(x) = 0$	sinon

Ces deux représentations d'un sous-ensemble sont différentes, indépendantes, mais équivalentes, en ce sens que l'on peut toujours en construire une à partir de l'autre. La plus prisee est évidemment la première. La raison est purement historique. On pourrait très bien utiliser la seconde et exprimer tous les concepts liés à la théorie des ensembles avec cette représentation. Définir un sous ensemble A d'un ensemble X revient à définir une fonction μ_A , dite d'appartenance ou à énumérer chacun des éléments y appartenant.

L'avantage de la seconde représentation est qu'elle nous permet de facilement définir la notion de sous-ensemble flou \tilde{A} d'un ensemble X , de manière cohérente avec le concept intellectuel qui lui correspond.

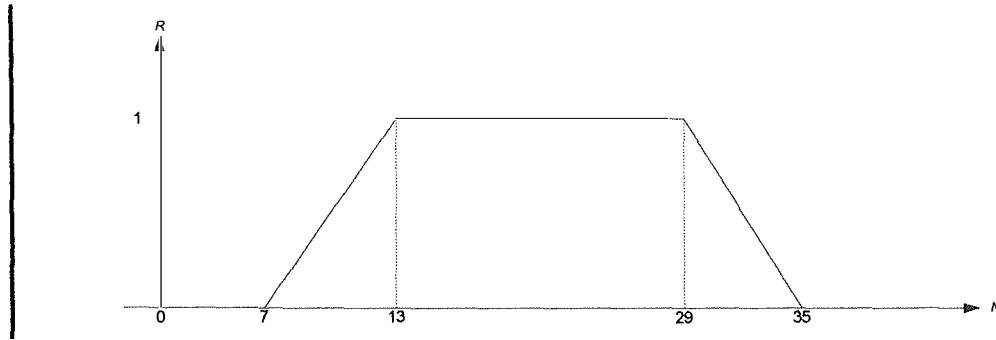
On définit **un sous-ensemble flou \tilde{A} d'un ensemble X** par une fonction $\mu_{\tilde{A}} : X \longrightarrow [0,1] \subseteq \mathbb{R}$, dite d'appartenance. L'énumération telle que celle vue précédemment, n'a plus de sens, à moins que pour chaque élément cité, on ne spécifie son degré d'appartenance à l'ensemble X . Un sous ensemble flou \tilde{A} d'un ensemble X ne se définit que par sa fonction d'appartenance.

Ex : Soit \mathbb{N} l'ensemble des naturels. Soit \tilde{A} « quelque chose » qui représente les âges admissibles pour un jeune être humain . On modélise \tilde{A} par un sous-ensemble flou \mathbb{N} défini par sa fonction d'appartenance $\mu_{\tilde{A}}$:

$$\mu_{\tilde{A}} : \mathbb{N} \longrightarrow \{0,1\} \subseteq \mathbb{N}$$

n	$\mu_{\tilde{A}}(n) = 0$	si $0 \leq n \leq 7$ ou $n \geq 35$
	$\mu_{\tilde{A}}(n) = (n-7)/6$	si $7 \leq n \leq 13$
	$\mu_{\tilde{A}}(n) = 1$	si $13 \leq n \leq 29$
	$\mu_{\tilde{A}}(n) = (35-n)/6$	si $29 \leq n \leq 35$

On considère ainsi que l'on est pas encore jeune avant 7 ans, que l'on est plus jeune après 35 ans, que l'on est jeune entre 13 et 29 ans, que notre jeunesse croît entre 7 et 13 ans et enfin qu'elle décroît entre 29 et 35 ans.



Remarquons que nous n'avons encore rien dit sur l'allure de cette fonction. A priori, il n'est pas nécessaire de limiter sa valeur à l'intervalle $[0,1]$. On pourrait imaginer des degrés d'appartenance négatifs ou supérieurs à 1. Cependant, quelle que soit la sémantique associée à ce degré d'appartenance, on suppose qu'il est toujours borné. On pourra toujours modifier, sans perte de généralité, cette sémantique de manière à le ramener dans l'intervalle $[0,1]$.

On définit $F(X)$ comme étant l'ensemble de tous les sous-ensembles flous de X .

Un sous-ensemble flou convexe \tilde{A} d'un ensemble X , est un sous-ensemble flou de X vérifiant la propriété suivante :

$$\forall x_1, x_2 \in X, \forall \lambda \in [0,1] \subseteq \mathbb{R}, \lambda \mu_{\tilde{A}}(x_1) + (1-\lambda)\mu_{\tilde{A}}(x_2) \leq \mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2)$$

Deux sous-ensembles flous \tilde{A}_1 et \tilde{A}_2 de X , sont dit égaux lorsque leur fonction d'appartenance respective, sont égales en tout point de X . On note cette propriété $\tilde{A}_1 = \tilde{A}_2$

3.3.2.2 Les α -coupes, le support et le support topologique d'un sous-ensemble flou

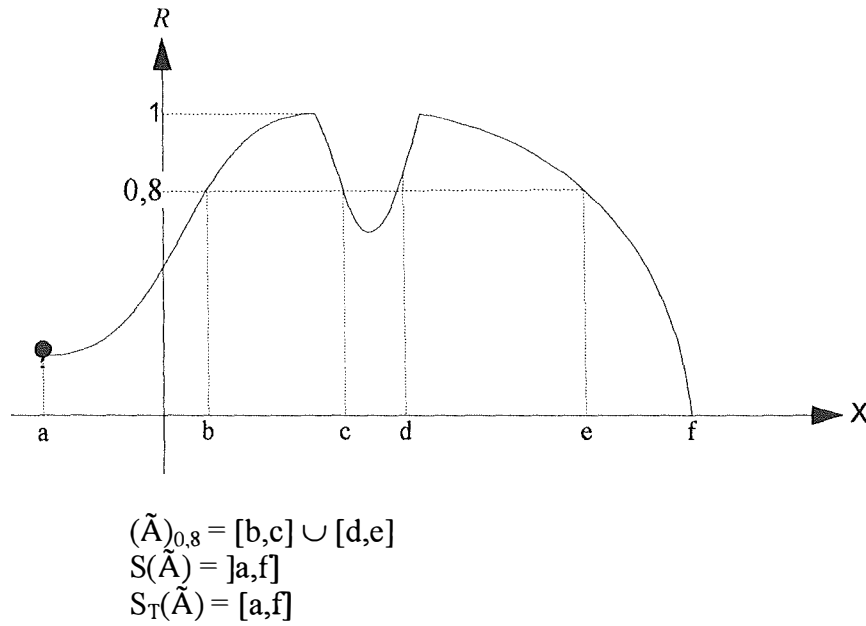
Soit $\tilde{A} \in F(X)$ de fonction d'appartenance $\mu_{\tilde{A}}(.)$. Pour tout α dans l'intervalle $[0,1]$, on définit l' α -coupes de \tilde{A} , notée $(\tilde{A})_\alpha$. On définit aussi le support de \tilde{A} , noté $S(\tilde{A})$, et le support topologique de \tilde{A} , noté $S_T(\tilde{A})$:

$$(\tilde{A})_\alpha = \{ x \in X \mid \mu_{\tilde{A}}(x) \geq \alpha \}$$

$$S(\tilde{A}) = \{ x \in X \mid \mu_{\tilde{A}}(x) \neq 0 \}$$

$$S_T(\tilde{A}) \text{ la fermeture de } S(\tilde{A})$$

Ex : Pour une fonction $\mu_{\tilde{A}}(.)$ d'appartenance particulière représentée géométriquement caractérisant un nombre flou \tilde{A} . Réalisons une coupe au niveau 0,8. Nous avons que



3.3.2.3 Les nombres flous

Un **nombre flou** \tilde{N} , est un sous-ensemble flou convexe de \mathfrak{R} vérifiant les propriétés suivantes :

$$\exists ! N \in \mathfrak{R} \text{ t.q. } \mu_{\tilde{N}}(N)=1$$

$\mu_{\tilde{N}}(.)$ est continue par morceaux sur \mathfrak{R}

N est appelé **valeur centrale** du nombre flou \tilde{N} . On appelle \tilde{N} un nombre flou, car il est fort pratique de le considérer comme une entité numérique valant « approximativement N ». Nous allons ainsi lui « appliquer » tous les opérateurs existant sur les nombres réels. L'approximation étant définie par la forme de la fonction d'appartenance. Si la fonction d'appartenance vaut 0 partout et 1 en N , alors il n'y a plus d'approximation et \tilde{N} peut être considéré comme le nombre réel N . Par convention nous noterons $\tilde{N} = N$ dans ce cas, et considérerons toujours que N est la valeur centrale de \tilde{N} .

En fonction de nos conventions, $F(\mathfrak{R})$ est donc **l'ensemble des nombres flous**. Nous notons aussi $F(\mathfrak{R})^+ \subseteq F(\mathfrak{R})$ l'ensemble des nombres flous dont la fonction d'appartenance est nulle pour toute valeur réelle négative. Cet ensemble caractérise en fait l'ensemble des valeurs « approximatives positives ».

Par définition on a évidemment les propriétés suivantes :

$$\mathfrak{R}^+ \subseteq \mathfrak{R} \subseteq F(\mathfrak{R})$$

$$\mathfrak{R}^+ \subseteq F(\mathfrak{R})^+ \subseteq F(\mathfrak{R})$$

On définit ainsi plusieurs types de nombres flous particuliers, en fonction de l'allure de leur fonction d'appartenance. Notamment, **on définit les nombres flous L-R**. L et R sont deux fonctions qui caractérisent la fonction d'appartenance du nombre flou, respectivement à gauche (L pour Left), et à droite (R pour Right) de la valeur centrale.

Soient

$$R : \mathfrak{R}^+ \longrightarrow [0,1] \subseteq \mathfrak{R} \quad \begin{array}{ll} \text{continue sur } \mathfrak{R}, \text{ strictement décroissante sur } [0,1]. \\ x \longmapsto \begin{array}{ll} R(x)=1 & \text{si } x=0 \\ 0 & \text{si } 1 \leq x \end{array} \end{array}$$

$$L : \mathfrak{R}^+ \longrightarrow [0,1] \subseteq \mathfrak{R} \quad \begin{array}{ll} \text{continue sur } \mathfrak{R}, \text{ strictement décroissante sur } [0,1]. \\ x \longmapsto \begin{array}{ll} L(x)=1 & \text{si } x=0 \\ 0 & \text{si } 1 \leq x \end{array} \end{array}$$

deux fonctions de référence (dites normalisées car elles varient dans $[0,1]$ seulement).

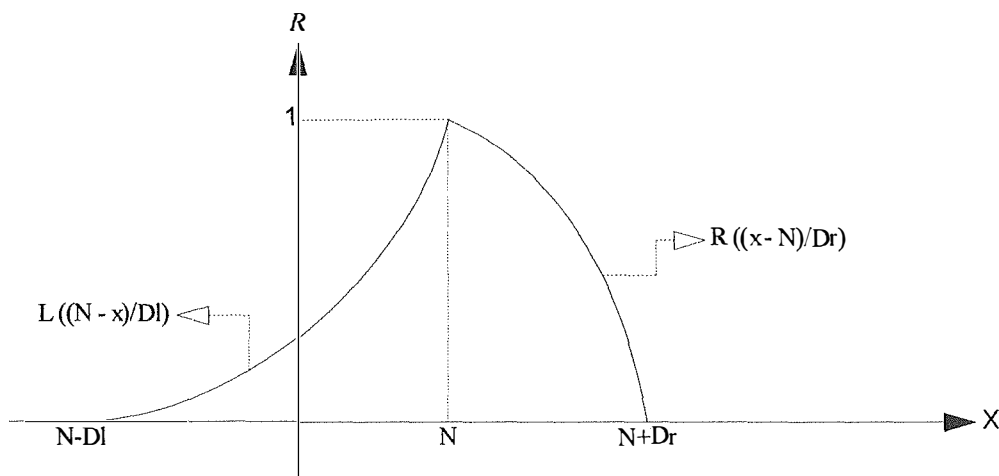
Un **nombre flou** \tilde{N} (de centre N) est dit de **type L-R**, si la propriété suivante est vérifiée :

$$\exists D_l, D_r \in \mathfrak{R}^+$$

$$\mu_{\tilde{N}}(x) = L((N-x)/D_l) \quad \text{si } x \leq N$$

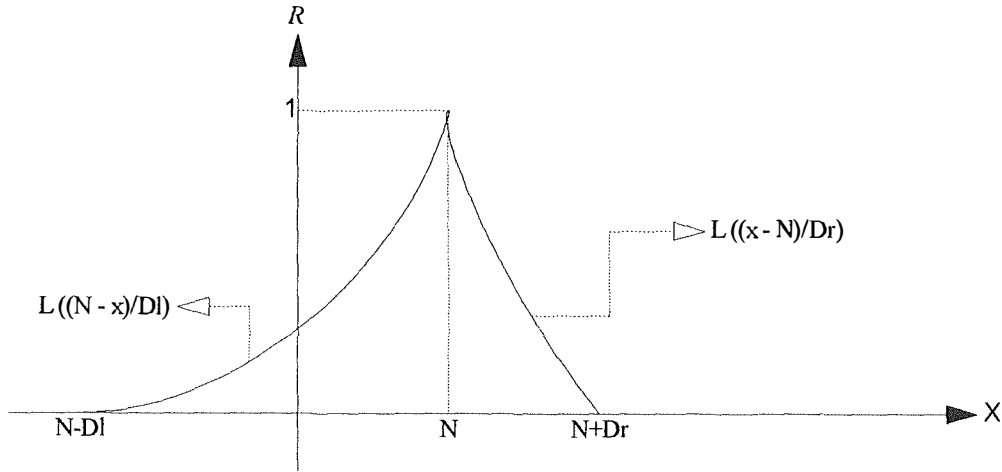
$$\mu_{\tilde{N}}(x) = R((x-N)/D_r) \quad \text{si } x \geq N$$

D_l et D_r sont respectivement appelés **dispersion gauche** et **dispersion droite** de \tilde{N} .



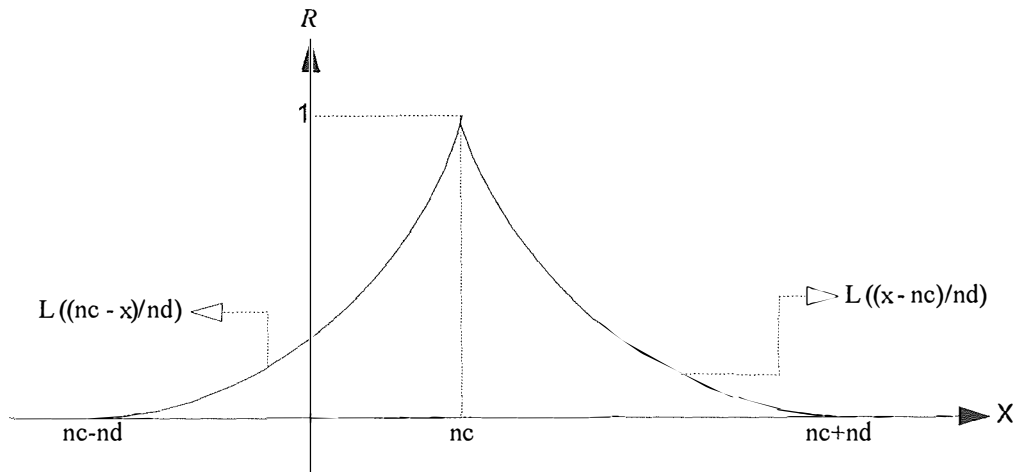
exemple de nombre flou de type L-R

Un **nombre flou** \tilde{N} de **type L-L**, est un nombre flou de type L-R où la fonction de référence R est définie de façon identique à la fonction L . Remarquons que cela ne signifie pas que la dispersion droite et la dispersion gauche de \tilde{N} soient égales.



exemple de nombre flou de type L-L

Un nombre \tilde{N} flou de type L-L, est dit **symétrique** lorsque sa dispersion droite et sa dispersion gauche sont égales. Un tel nombre flou n'est plus caractérisé que par trois paramètres : une valeur centrale N, une dispersion D et une fonction repère L(.). Remarquons que le qualificatif symétrique n'est défini que pour les nombres flous de type L-L. L'ensemble des nombres flous de type L-L est noté $F_{L-L}(\mathfrak{R})$ pour une fonction repère L(.).



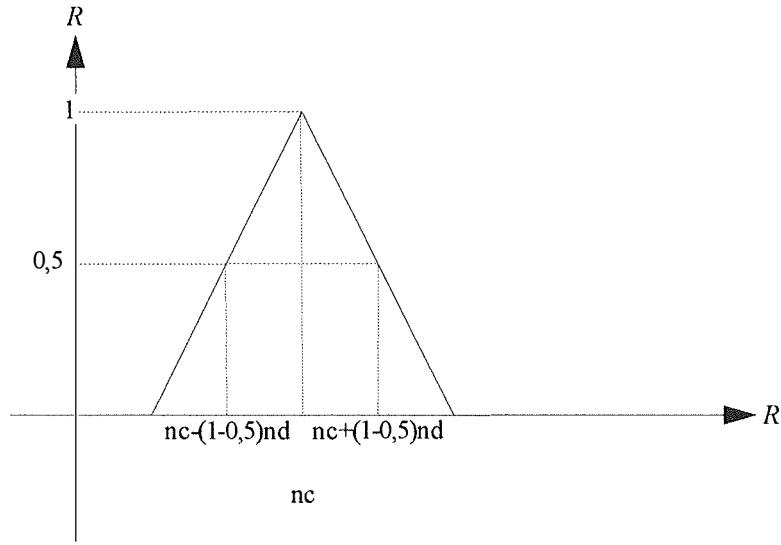
exemple de nombre flou de type L-L symétrique

Soit $\tilde{n} \in F_{L-L}(\mathfrak{R})$ symétrique. Par convention, on le note $(n_c, n_d)_L$, n_c étant la valeur (réelle) centrale de N (dans ce cas égale à N), n_d sa dispersion (réel), et L(.) sa fonction repère permettant de définir la forme du nombre.

La fonction L n'est normalement pas inversible car pour toute valeur plus grande que 1, son image par L vaut 0 par définition. On peut cependant se limiter l'intervalle $[0,1]$, sans perdre de généralité. Dans ce cas L devient inversible. Posons que $L^{-1}(1) = 0$ et notons $L^{-1}(\cdot)$ l'inverse de L(.) sur l'intervalle $[0,1]$. On montre ainsi aisément que

$$\forall \alpha \in [0,1] \subseteq \mathfrak{R}, \forall \tilde{n} \in F_{L-L}(\mathfrak{R}) \text{ et symétrique} \\ ((n_c, n_d)_L)_\alpha = [n_c - L^{-1}(\alpha) n_d, n_c + L^{-1}(\alpha) n_d] \subseteq \mathfrak{R}$$

Ex : $L(x) = \text{Max}(1-x, 0)$ définie sur \mathfrak{R}^+
 $\forall x \in [0,1] \subseteq \mathfrak{R}, L^{-1}(x) = 1-x = L(x)$
 Considérons un $(n_c, n_d)L \in F_{L-L}(\mathfrak{R})$ et symétrique



3.3.2.4 Le principe d'extension en général

Ce principe a pour objectif d'étendre toute fonction d'un ensemble A , ou tout produit cartésien d'ensembles $A_1 \times A_2$ à valeurs dans un ensemble A_3 , à une fonction de même signature où les ensembles sont remplacés par des sous-ensembles flous de ces ensembles. Mathématiquement, considérons n'importe quelle fonction de la forme :

$$\left. \begin{array}{l} f: A_1 \longrightarrow A_3 \\ a_1 \longrightarrow a_3 = f(a_1) \end{array} \right| \begin{array}{l} f: A_1 \times A_2 \longrightarrow A_3 \\ (a_1, a_2) \longrightarrow a_3 = f(a_1, a_2) \end{array}$$

Ce principe a pour but d'étendre f , de telle manière que sa signature devienne

$$\left. \begin{array}{l} f: F(A_1) \longrightarrow F(A_3) \\ (\tilde{a}_1) \longrightarrow \tilde{a}_3 = f(\tilde{a}_2) \end{array} \right| \begin{array}{l} f: F(A_1) \times F(A_2) \longrightarrow F(A_3) \\ (\tilde{a}_1, \tilde{a}_2) \longrightarrow \tilde{a}_3 = f(\tilde{a}_1, \tilde{a}_2) \end{array}$$

et que l'on ait au moins la propriété suivante :

$$\mu_{\tilde{a}_1}(a_1) = 1 \text{ et } \mu_{\tilde{a}_2}(a_2) = 1 \Rightarrow \mu_{\tilde{a}_3}(f(a_1, a_2)) = 1$$

et plus particulièrement

$$\tilde{a}_1 = a_1 \text{ et } \tilde{a}_2 = a_2 \Rightarrow c = f(\tilde{a}_1, \tilde{a}_2) = f(a_1, a_2) = a_3$$

Ex : Considérons la fonction et loi interne d'addition des réels

$$+ : \mathfrak{R} \times \mathfrak{R} \longrightarrow \mathfrak{R} \\ (a_1, a_2) \quad a_3 = +(a_1, a_2) \text{ que l'on note } a_1 + a_2$$

ce principe a pour but d'étendre l'addition de nombres réels à l'addition de nombres flous. Exactement comme on l'étend à l'addition des nombre complexes, ou des variables aléatoires.

$$+ : F(\mathfrak{R}) \times F(\mathfrak{R}) \longrightarrow F(\mathfrak{R}) \\ (\tilde{a}_1, \tilde{a}_2) \quad \tilde{a}_3 = +(\tilde{a}_1, \tilde{a}_2) \text{ que l'on note } \tilde{a}_1 + \tilde{a}_2$$

de telle manière que si $\tilde{a}_1 = a_1$ et $\tilde{a}_2 = a_2$ alors $\tilde{a}_3 = \tilde{a}_1 + \tilde{a}_2 = a_1 + a_2 = a_3$, ce qui revient à dire que l'addition de nombre flous qui sont des réels, n'est autre que l'addition de ces nombres réels.

Remarquons que nous n'avons encore que l'objectif, les spécifications, du principe d'extension, mais pas encore de manière de les réaliser. Pour cela il faudrait définir la fonction d'appartenance $\mu_{\tilde{a}_3}(\cdot)$ à partir des deux fonctions d'appartenance $\mu_{\tilde{a}_1}(\cdot)$ et $\mu_{\tilde{a}_2}(\cdot)$ lorsque $\tilde{a}_3 = f(\tilde{a}_1, \tilde{a}_2)$. Il existe plusieurs façons de procéder en fonction de la sémantique que l'on accorde au degré d'appartenance. Ces différentes manières de procéder représentent en fait les différents choix d'extension dont nous parlons dans l'introduction de ce chapitre.

3.4 La théorie possibiliste des ensembles flous

3.4.1 La justification possibiliste du principe d'extension de Zadeh, Dubois et Prade

Zadeh, Dubois et Prade font le choix de considérer la théorie des ensembles flous dans le cadre d'une interprétation possibiliste du degré d'appartenance. Rappelons que jusqu'à présent, la théorie des ensembles flous, basée sur la notion de degré d'appartenance, et celle des possibilités, basée sur la notion de possibilité, avaient été présentées indépendamment et sans lien. Théoriquement cette démarche est tout à fait correcte. Lorsque l'on lie le degré d'appartenance et la possibilité, le principe d'extension prend une direction particulière. Nous expliquons cette direction, dans le cas d'une fonction à deux variables. Le cas d'une fonction à une variable est, à notre avis, directement compréhensible à partir de celui-ci.

Considérons une variable a_1 (resp. a_2, a_3) non probabiliste, à valeur dans A_1 (resp. A_2, A_3) un ensemble fini, et \tilde{a}_1 (resp. \tilde{a}_2, \tilde{a}_3) un sous ensemble flou de A_1 (resp. A_2, A_3).

Donnons à $\mu_{\tilde{a}_1}(\cdot)$ (resp. $\mu_{\tilde{a}_2}(\cdot)$) la sémantique d'une distribution de possibilité de a_1 (resp. a_2) dans A_1 (resp. A_2).

Autrement dit on pose que

$$\forall x \in A_1, \mu_{\tilde{a}_1}(x) = \text{la possibilité que la variable } a_1 \text{ prenne la valeur } x \quad (P(a_1 = x))$$

$\forall y \in A_2, \mu_{a_2}(y) = \text{la possibilité que la variable } a_2 \text{ prenne la valeur } y \quad (P(a_2 = y))$

Considérons une fonction f dont la signature est

$$f: A_1 \times A_2 \longrightarrow A_3$$

$$(a_1, a_2) \quad a_3 = f(a_1, a_2)$$

que l'on veut étendre à une fonction f dont la signature deviendra

$$f: F(A_1) \times F(A_2) \longrightarrow F(A_3)$$

$$(\tilde{a}_1, \tilde{a}_2) \quad \tilde{a}_3 = f(\tilde{a}_1, \tilde{a}_2)$$

de telle manière que la sémantique possibiliste associée aux fonctions d'appartenance soit conservée. Autrement dit, supposons que $\tilde{a}_3 = f(\tilde{a}_1, \tilde{a}_2)$. On désire que :

$\forall z \in A_3 \quad \mu_{\tilde{a}_3}(z) = \text{la possibilité que la variable } a_3 \text{ prenne la valeur } z \quad (P(a_3 = z))$

Pour cela notons $f^{-1}(z) \equiv \{(x_1, y_1), \dots, (x_n, y_n)\}$. Par définition, cet ensemble est soit vide, soit fini. On désire donc que :

$$\mu_{\tilde{a}_3}(z) = P(a_3 = z)$$

- Si $f^{-1}(z)$ est vide alors l'événement ayant la propriété $(a_3 = z)$ est impossible. Il s'en suit donc que $P(a_3 = z) = 0$, par définition d'une possibilité.
- Sinon, on a que l'événement ayant la propriété $(a_3 = z)$ se réalise dans un des n cas. Nous pouvons l'énoncer de la manière suivante

$$P(a_3 = z) = P(((a_1, a_2) \equiv (x_1, y_1)) \text{ ou } \dots \text{ ou } ((a_1, a_2) \equiv (x_n, y_n)))$$

$$= P((a_1 = x_1 \text{ et } a_2 = y_1) \text{ ou } \dots \text{ ou } (a_1 = x_n \text{ et } a_2 = y_n))$$

Valeur que l'on peut calculer selon les lois de la théorie des possibilités

$$= \text{Sup} \{ P(a_1 = x_1 \text{ et } a_2 = y_1), \dots, P(a_1 = x_n \text{ et } a_2 = y_n) \}$$

$$= \text{Sup} \{ \text{Min}(\text{Pos}(a_1 = x_1), \text{Pos}(a_2 = y_1)), \dots, \text{Min}(\text{Pos}(a_1 = x_n), \text{Pos}(a_2 = y_n)) \}$$

Expression que l'on peut exprimer à partir des fonctions d'appartenance $\mu_{a_1}(x)$ et $\mu_{a_2}(y)$

$$= \text{Sup} \{ \text{Min}(\mu_{a_1}(x_1), \mu_{a_2}(y_1)), \dots, \text{Min}(\mu_{a_1}(x_n), \mu_{a_2}(y_n)) \}$$

$$= \text{Sup} \{ \text{Min}(\mu_{a_1}(x), \mu_{a_2}(y)) \mid (x, y) \in f^{-1}(z) \}$$

3.4.2 Le principe d'extension proposé par Zadeh, Dubois et Prade

On comprend maintenant aisément le principe d'extension particulier énoncé par Z.D.H.

$$\begin{array}{l|l}
 f: A_1 \longrightarrow A_3 & f: A_1 \times A_2 \longrightarrow A_3 \\
 a_1 \quad a_3 = f(a_1) & (a_1, a_2) \quad a_3 = f(a_1, a_2) \\
 \forall z \in \mathfrak{R} \text{ notons} & \forall z \in \mathfrak{R} \text{ notons} \\
 f^{-1}(z) = \{ x \in \mathfrak{R} \mid f(x) = z \} & f^{-1}(z) = \{ (x, y) \in \mathfrak{R} \times \mathfrak{R} \mid f(x, y) = z \}
 \end{array}$$

La signature de f étendue devient donc

$$\begin{array}{l|l}
 f: F(A_1) \longrightarrow F(A_3) & f: F(A_1) \times F(A_2) \longrightarrow F(A_3) \\
 \tilde{a}_1 \quad \tilde{a}_3 = f(\tilde{a}_1) & (\tilde{a}_1, \tilde{a}_2) \quad \tilde{a}_3 = f(\tilde{a}_1, \tilde{a}_2)
 \end{array}$$

Dans le cas d'une fonction à une variable, la fonction d'appartenance $\mu_{\tilde{a}_3}(\cdot)$ de l'image de \tilde{a}_1 par f étendue est définie à partir de $\mu_{\tilde{a}_1}(\cdot)$ comme suit

$$\boxed{\forall z \in \mathfrak{R} \quad \mu_{\tilde{a}_3}(z) = \begin{cases} \sup \{ \mu_{\tilde{a}_1}(x) \mid x \in f^{-1}(z) \} & \text{si } f^{-1}(z) \neq \emptyset \\ 0 & \text{sinon} \end{cases}}$$

Dans le cas d'une fonction à deux variables, la fonction d'appartenance $\mu_{\tilde{a}_3}(\cdot)$ de l'image de \tilde{a}_1 et \tilde{a}_2 par f étendue est définie à partir de $\mu_{\tilde{a}_1}(\cdot)$ et $\mu_{\tilde{a}_2}(\cdot)$ comme suit

$$\boxed{\forall z \in \mathfrak{R} \quad \mu_{\tilde{a}_3}(z) = \begin{cases} \sup \{ \min(\mu_{\tilde{a}_1}(x), \mu_{\tilde{a}_2}(y)) \mid (x, y) \in f^{-1}(z) \} & \text{si } f^{-1}(z) \neq \emptyset \\ 0 & \text{sinon} \end{cases}}$$

Le cas de fonctions à plus de deux variables est déductible de celui à deux variables.

Ex : Considérons les nombres flous 2 et 6 dont les fonctions d'appartenance sont

$$\forall x \in \mathfrak{R} \quad \mu_2(x) = \begin{cases} 1 & \text{si } x = 2 \\ 1/2 & \text{si } x = 3 \\ 0 & \text{dans les autres cas} \end{cases}$$

$$\forall x \in \mathfrak{R} \quad \mu_6(x) = \begin{cases} 1 & \text{si } x = 6 \\ 1/2 & \text{si } x = 5 \\ 1/4 & \text{si } x = 3 \\ 0 & \text{dans les autres cas} \end{cases}$$

Si l'on étend la multiplication \times de nombres réels à la multiplications \times de nombres flous suivant le principe d'extension proposé par Z.D.P., alors le nombre flou 2×6 obtenu par multiplication de 2 et 6 est caractérisé par une fonction d'appartenance $\mu_{2 \times 6}(\cdot)$ définie comme suit :

$$\forall x \in \mathfrak{R} \quad \mu_{2 \times 6}(x) = \begin{cases} 1 & \text{si } x = 12 \\ 1/2 & \text{si } x \in \{10, 15, 18\} \subseteq \mathfrak{R} \\ 1/4 & \text{si } x \in \{14, 21\} \subseteq \mathfrak{R} \\ 0 & \text{dans les autres cas} \end{cases}$$

3.4.3 Une propriété intéressante de Hung T. Nguyen

Dans son article [Nguyen78], Hung T. Nguyen étudie plus en profondeur les différentes propriétés découlant du principe d'extension de Z.D.P.. Notamment, il nous donne une condition nécessaire et suffisante, très utile pour le calcul des α -coupes. La propriété est la suivante :

Si $f : F(A_1) \times F(A_2) \longrightarrow F(A_3)$, définie par extension de $f : A_1 \times A_2 \longrightarrow A_3$ selon le principe d'extension de Z.D.P., alors

$$\begin{array}{l} \forall \tilde{a}_1 \in F(A_1), \forall \tilde{a}_2 \in F(A_2) \text{ et } \mu_{\tilde{a}_1}(\cdot) \text{ et } \mu_{\tilde{a}_2}(\cdot) \text{ leurs fonction d'appartenance respectives} \\ \forall \alpha \in]0,1] \subseteq \mathfrak{R} \\ (f(\tilde{a}_1, \tilde{a}_2))_\alpha = f((\tilde{a}_1)_\alpha, (\tilde{a}_2)_\alpha) = \{z \in A_3 \mid z = f(x,y) \ \forall x \in (\tilde{a}_1)_\alpha, \forall y \in (\tilde{a}_2)_\alpha\} \\ \Leftrightarrow \\ \forall z \in A_3, \text{ le Sup } \{ \text{Min}(\mu_{\tilde{a}_1}(x), \mu_{\tilde{a}_2}(y)) \mid (x,y) \in f^{-1}(z) \} \text{ est atteint} \end{array}$$

Même si l'on ne peut pas, de manière générale, « calculer facilement » l'image de deux ensembles flous par une fonction étendue, on peut, dans certains cas, « calculer beaucoup plus facilement » les α -coupes de cette image. Nous verrons un peu plus loin une application de cette propriété.

3.4.4 Trois indices de comparaison de nombres flous proposés par Dubois et Prade

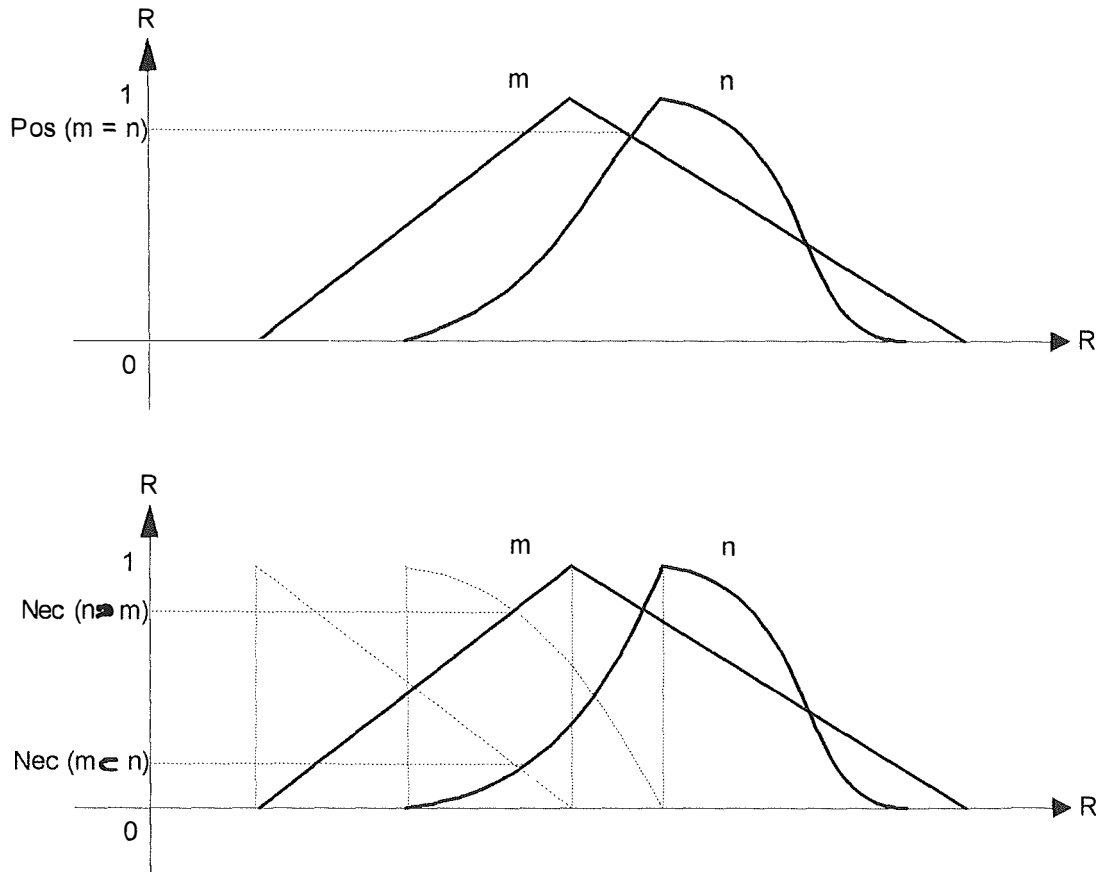
Comme nous le verrons par la suite, il est utile de pouvoir comparer deux nombres flous. Dubois et Prade proposent trois indices pour cela. Le premier représente ce que l'on pourrait appeler un degré d'égalité, les deux autres ce que l'on pourrait appeler un degré d'inclusion de deux nombres flous.

Leur définition mathématique est la suivante. Soit $m, n \in F(\mathfrak{R})$ caractérisés par leur fonction d'appartenance respective $\mu_m(\cdot)$ et $\mu_n(\cdot)$.

$$\begin{array}{l} \text{Pos}(m = n) = \text{Sup } \{ \text{Min}(\mu_m(x), \mu_n(x)) \mid x \in \mathfrak{R} \} \\ \text{Nes}(m \subset n) = \text{Inf } \{ \text{Max}(1 - \mu_m(x), \mu_n(x)) \mid x \in \mathfrak{R} \} \\ \text{Nes}(m \supset n) = \text{Inf } \{ \text{Max}(\mu_m(x), 1 - \mu_n(x)) \mid x \in \mathfrak{R} \} \end{array}$$

A cause de la sémantique que l'on donne au degré d'appartenance dans la théorie possibiliste des ensembles flous, le premier indice est considéré comme une mesure de possibilité et le second comme une mesure de nécessité, d'où les notations.

Nous montrons géométriquement la signification de ces trois indices sur les deux schéma qui suivent.



Les propriétés suivantes sont alors facilement démontrables

$$\forall x \in \mathfrak{R} \mu_m(x) = \mu_n(x) \text{ (« } m = n \text{ »)} \Rightarrow \text{Pos} (m = n) = 1$$

$$\forall x \in \mathfrak{R} \mu_m(x) \leq \mu_n(x) \text{ (« } m \text{ inclus dans } n \text{ »)} \Rightarrow \text{Nec} (m \subset n) = 1$$

$$\forall x \in \mathfrak{R} \mu_m(x) \geq \mu_n(x) \text{ (« } n \text{ inclus dans } m \text{ »)} \Rightarrow \text{Nec} (n \supset m) = 1$$

Remarquons que l'on a pas les équivalences dans $F(\mathfrak{R})$, mais bien dans \mathfrak{R} . Nous laissons au lecteur le soin d'étudier la manière dont se comporte ces indices dans des cas particuliers. Par exemple, que se passe-t-il lorsqu'un des deux nombres est purement flou et l'autre est un nombre réel ? c'est exactement la situation à laquelle nous serons confrontés lors de l'élaboration de notre solution à un des trois problèmes de production.

3.4.5 Le calcul de la multiplication et de l'addition de nombres flous

La définition de l'extension d'une fonction sur les réels selon le principe d'extension de Z.D.P. n'est pas pratique, pour le calcul proprement dit de l'image de deux nombres flous par cette fonction étendue. Ceci exactement comme la définition d'une dérivée en un point d'une fonction, n'est pas pratique pour le calcul proprement dit de la dérivée d'une fonction quelconque en un point. Pour certains types de fonctions, dites dérivables, ce calcul est fortement simplifié. Il existe des règles facilitant le calcul. De même, dans certains cas

particuliers de nombres flous, ce calcul est rendu beaucoup plus facile. Par exemple dans la cas de nombres flous vérifiant la propriété de Nguyen.

Soit $+$ et \times respectivement les loi d'addition et de multiplication des nombres flous en général, définies par extension de l'addition et de la multiplication dans les réels selon le principe d'extension de Z.D.P. Soit L une fonction repère. La propriété de Nguyen est évidemment vérifiée pour les nombres flous de type L-L, car leur fonction d'appartenance est par définition continue sur \mathfrak{R} . En l'appliquant, on comprend la propriété qui en découle :

Soit $m, n \in F_{L-L}(\mathfrak{R})^+$. On note aussi $m = (m_c, m_d)_L$ et $n = (n_c, n_d)_L$. Alors

$$\forall \alpha \in]0,1] \subseteq \mathfrak{R}$$

$$\begin{aligned} (m + n)_\alpha &= ((m_c, m_d)_L + (n_c, n_d)_L)_\alpha \\ &= ((m_c, m_d)_L)_\alpha + ((n_c, n_d)_L)_\alpha \\ &= [m_c - L^{-1}(\alpha) m_d, m_c + L^{-1}(\alpha) m_d] + [n_c - L^{-1}(\alpha) n_d, n_c + L^{-1}(\alpha) n_d] \\ &= [(m_c - L^{-1}(\alpha) m_d) + (n_c - L^{-1}(\alpha) n_d), (m_c + L^{-1}(\alpha) m_d) + (n_c + L^{-1}(\alpha) n_d)] \end{aligned}$$

$$\begin{aligned} (m \times n)_\alpha &= ((m_c, m_d)_L \times (n_c, n_d)_L)_\alpha \\ &= ((m_c, m_d)_L)_\alpha \times ((n_c, n_d)_L)_\alpha \\ &= [m_c - L^{-1}(\alpha) m_d, m_c + L^{-1}(\alpha) m_d] \times [n_c - L^{-1}(\alpha) n_d, n_c + L^{-1}(\alpha) n_d] \\ &= [(m_c - L^{-1}(\alpha) m_d) \times (n_c - L^{-1}(\alpha) n_d), (m_c + L^{-1}(\alpha) m_d) \times (n_c + L^{-1}(\alpha) n_d)] \end{aligned}$$

On remarque que l'addition de deux nombres flous de type L-L est une loi interne. En effet, si l'on restreint $L(\cdot)$ à l'intervalle $[0,1]$, L devient inversible et on peut généraliser la relation ci-dessus :

$$\forall \alpha \in [0,1] \subseteq \mathfrak{R}$$

$$\begin{aligned} (m + n)_\alpha &= [(m_c - L^{-1}(\alpha) m_d) + (n_c - L^{-1}(\alpha) n_d), (m_c + L^{-1}(\alpha) m_d) + (n_c + L^{-1}(\alpha) n_d)] \\ &= [(m_c + n_c) - L^{-1}(\alpha) (m_d + n_d), (m_c + n_c) + L^{-1}(\alpha) (m_d + n_d)] \end{aligned}$$

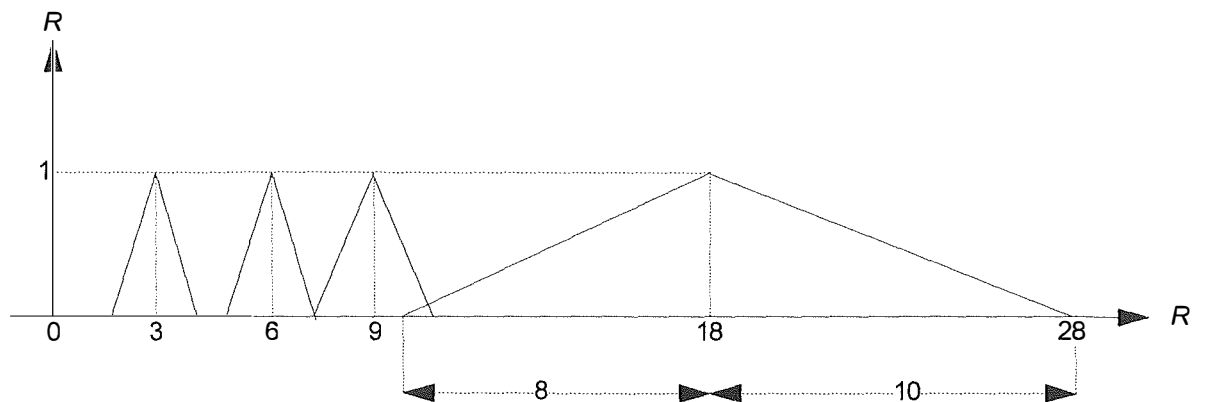
ce qui est équivalent à dire

$$m + n = (m_c, m_d)_L + (n_c, n_d)_L = (m_c + n_c, m_d + n_d)_L$$

Il n'en est pas de même pour la multiplication.

Ex1 : Soit $L(x) = \text{Max}(1-x, 0)$ une fonction de référence

Soit $(3,1)_L, (6,1)_L \in F_{L-L}(\mathfrak{R})^+$, Nous montrons géométriquement ce que donne leur multiplication et leur addition.



$$(3,1)_L + (6,1)_L = (9,2)_L \in F_{L-L}(\mathfrak{R})^+$$

$(3,1)_L \times (6,1)_L \in F(\mathfrak{R})^+$ un nombre flou dont le centre vaut 18, la dispersion gauche vaut 8, tandis que la droite vaut 10.

Chapitre 4

**A la recherche d'une solution « floue » à
un des trois problèmes**

4. A la recherche d'une solution « floue » à un des trois problèmes

4.1 Que chercher ...

Lorsque cette partie du stage est abordée, notre connaissance sur la théorie des ensembles flous est encore fort floue. Nous avons déjà rencontré beaucoup de concepts liés à cette théorie, mais nous sommes encore loin de les maîtriser, de les « sentir », et d'imaginer l'intérêt que cette « nouvelle » théorie suscite dans le monde scientifique et industriel. Nous prions donc le lecteur de se mettre à notre place à ce moment-là, et de se comporter comme s'il n'avait pas encore lu dans les détails le chapitre d'introduction à la théorie des ensembles flous. Nous voilà donc face à une jungle, dans laquelle il nous faut trouver le chemin d'une éventuelle « solution » à l'un des trois problèmes posés. Nous appelons solution à l'un des trois problèmes, une technique mathématique qui, d'une part répond aux « spécifications » de leur énoncé donné dans la partie de présentation de l'ARBED, et d'autre part, passe outre des insuffisances révélées par les solutions actuelles. « Que chercher » est alors la première question qui nous vient à l'esprit. Pour chacun des trois problèmes, expliquons maintenant la manière dont nous avons essayé d'y répondre.

4.1.1 ... pour le problème direct ?

Nous remarquons que les insuffisances de la « solution » actuelle à ce problème sont liées au fait que l'on ne connaît qu'approximativement les teneurs de chaque type de mitraille en chaque élément. Nous avons là un problème typiquement dans le domaine couvert par la théorie des ensembles flous. Après tout, son but est bien de représenter et manipuler des connaissances approximatives ! Notre idée est donc de reprendre la solution actuelle et de voir si elle n'a jamais été généralisée avec des concepts flous. N'existe-t-il pas, quelque part dans la jungle de la théorie des ensembles flous, quelqu'un qui s'est un jour intéressé à la résolution d'un PL, mais où les variables et constantes ne sont plus des réels mais des concepts flous (exactement comme on pourrait les remplacer par des concepts probabilistes) ? Des nombres flous par exemple ... C'est là, la direction dans laquelle nous chercherons. L'énoncé flou du problème direct aurait donc la forme suivante :

4.1.1.1 Les hypothèses

$\forall m \in \mathcal{M} \quad \forall e \in \mathcal{E},$	$T(m,e)$ un concept flou et F l'espace correspondant.
$\forall e \in \mathcal{E} \quad \forall c \in \mathcal{C},$	$a(e,c)$ un concept flou et F l'espace correspondant.
$\forall m \in \mathcal{M} \quad \forall c \in \mathcal{C},$	$q(c,m)$ un concept flou et F l'espace correspondant.
\times	$=$ une loi interne de « multiplication » sur F
\sum	$=$ une sommation selon une loi interne « d'addition » sur F .
MIN	$=$ une minimisation selon une « relation d'ordre » dans F .
\approx	$=$ une relation « d'égalité » dans F .

<p>Les C charges n'ont pas encore été produites. Il s'en suit donc que</p> <p>$\forall m \in \mathcal{M} \forall e \in \mathcal{E}$, $t(m,e)$ est connue approximativement à partir de la carte d'identité de la mitraille</p> <p>$\forall e \in \mathcal{E} \forall c \in \mathcal{C}$, $a(e,c)$ est connue. C'est l'analyse désirée.</p> <p>$\forall m \in \mathcal{M} \forall c \in \mathcal{C}$, $q(c,m)$ est inconnue</p>
--

Remarquons que dans notre cas, seul $t(m,e)$ doit nécessairement être flou. Les hypothèses posées ici sont donc plus générales.

4.1.1.2 Un énoncé flou du problème direct

Suite aux nouvelles hypothèses posées, une version fuzzyfiée du problème inverse aurait la forme suivante :

$\forall c \in \mathcal{C}$, trouver la solution du « PL flou » suivant :

$$MIN \sum_{m \in \mathcal{M}} ct(m) \times q(c,m)$$

$$SC \sum_{m \in \mathcal{M}} t(m,e) \times q(c,m) \approx a(e,c) \quad \forall e \in \mathcal{E}$$

$$\sum_{m \in \mathcal{M}} q(c,m) \approx U_c$$

« Pour chaque charge à produire, trouver les quantités approximatives de chacun des types de mitraille à enfourner, de manière à ce que la qualité des charges produites soit la plus proche possible de la qualité demandée, et que les coûts soient les plus petits possibles ».

4.1.2 ... pour le problème inverse ?

L'idée ici suit le même raisonnement que pour le problème direct. En effet, dans ce cas, l'insuffisance de la solution actuelle est liée au fait que les quantités utilisées de chaque type de mitraille ne sont connues qu'approximativement. De même, les teneurs de chaque type de mitraille n'étant pas constantes en tant que réels, on peut imaginer qu'elles le soient approximativement, c'est à dire de manière floue. Nous chercherons une technique de régression linéaire multiobjectifs, où les coefficients et variables seront des concepts flous et plus des réels. L'énoncé flou au problème direct aurait donc la forme suivante

4.1.2.1 Conventions supplémentaires de notation

A cause des insuffisances de la solution du problème direct, introduisons des conventions supplémentaires. Celles-ci viennent simplement s'ajouter à celles établies dans

le chapitre sur la modélisation des trois problèmes. Nous les avons mises à part, car elles sont propres à ce problème.

F	= un espace lié à la théorie des sous-ensembles flous.
\otimes	= « une » opération de « multiplication » de signature : $F^{C \times M} \times F^M \rightarrow F^C$
\approx	= une relation d'égalité dans F^C

4.1.2.2 Les hypothèses

Les C charges ont été produites

- $\forall m \in M \quad \forall c \in C, q(m,c)$ est un concept flou. Soit F l'espace auquel cette valeur appartient ($\Rightarrow Q \in F^{C \times M}$). Cette valeur est connue et est la quantité approximative utilisée (réellement) lors de la production, déterminée avec la solution du problème direct (bien que l'on savait que la solution trouvée ne serait pas exactement ce que l'on voulait).
- $\forall e \in E \quad \forall c \in C, a(e,c)$ est un concept flou. Soit F l'espace auquel cette valeur appartient ($\Rightarrow A_e \in F^C$). Cette valeur est connue et est la quantité mesurée (réellement) lors de l'analyse de la charge produite. Rappelons qu'elle est très probablement différente de celle désirée.
- On considère la teneur de chaque type de mitraille en l'élément e . Cette valeur n'a pas de sens en tant que réel. Nous utilisons donc un concept flou. Soit F l'espace auquel appartient cette valeur.
 $\forall m \in M \quad \forall e \in E, t(m,e) \in F (\Rightarrow T_e \in F^M)$ et est inconnue

4.1.2.3 Un énoncé flou du problème inverse

$\forall e \in E$, déterminer T_e TQ $Q \otimes T_e \approx A_e$

« Pour chaque élément, trouver la teneur floue de chacun de types de mitraille en chaque élément, connaissant approximativement la quantité de chaque type de mitraille utilisée pour chaque charge, la composition exacte de chaque charge en chaque élément et la relation qui les lie par l'intermédiaire de cette teneur »)

Lorsque l'on parle de régression linéaire comme ci-dessus, la matrice Q est appelée matrice des variables d'entrées, tandis que A_e est appelé vecteur des variables de sortie.

Si la résolution d'un tel système n'a pas de solution exacte dans les réels, on peut imaginer qu'il n'en est pas de même en flou. Par similarité avec son énoncé probabiliste, cet énoncé flou du problème inverse est très probablement insoluble de manière générale. Cependant, on peut imaginer qu'il devienne soluble en « coupant » la difficulté, comme nous l'avons fait pour trouver une solution au problème inverse selon son énoncé probabiliste. L'idée est alors de ne plus considérer des intervalles de confiance à $x\%$, mais plutôt des intervalles de possibilité à « x degrés de possibilité ». Les α -coupes par exemple ...

4.1.3 ... pour le problème de planification de production ?

Lors de notre première rencontre avec Monsieur Lescrenier, celui-ci nous fournit d'emblée de la documentation sur une solution floue à ce problème. Il s'agit en fait d'articles écrits en allemand, par W. Limbeck et H.-P. Lipp. Il semble qu'ils ont développé un système d'aide à la décision, basé sur la théorie des ensembles flous pour résoudre le problème de planification, en collaboration avec l'usine métallurgique Krupp Mannesman sprl et MIT. Ce problème est fort similaire à celui traité par profilARBED. Ne parlant pas l'allemand, avec l'aide d'une amie traductrice (merci Laurence), nous avons vite compris que leur solution était basée sur le concept de réseau de Petri flou. C'est donc dans cette direction que nous avons décidé de chercher.

4.2 Où chercher ?

Maintenant que nous avons assez précisément déterminé ce que nous devons chercher, il nous faut maintenant déterminer où il faut chercher. Ceci d'autant plus vite que notre temps est compté. Il faut que dans les trois mois, nous ayons débouché sur une piste floue intéressante à propos d'un des trois problèmes. Dans la « brousse » de la théorie des ensembles flous, des sciences en général et de l'industrie sidérurgique, connaissant la description de l'endroit où nous devons aller, nous devons demander notre chemin aux autochtones, de manière à y arriver en un minimum de temps. Nous avons là un beau challenge, typiquement dans la lignée de ce que l'on rencontre dans les entreprises privées. Nous décidons de les aborder de trois manières différentes.

La première est une tentative à court terme. On essaye d'exploiter ce « formidable » container à information que constitue le réseau public InterNet, auquel nous avons la chance d'avoir accès aux FUNDP. Nous essayerons de trouver des sites spécialisés dans le domaine dans le but de rapatrier de l'information. Nous essayerons aussi de trouver des NewsGroup intéressantes encore de trouver des adresses IP de spécialistes. Nous pourrions ainsi les contacter par courrier électronique. Nous avons là de quoi nous servir ... On trouve sur InterNet beaucoup de scientifiques et de chercheurs. C'est normal quand on sait que ce réseau s'est notamment développé à partir de la connexion d'universités (remplies de chercheurs avides de connaissances).

La deuxième manière est plutôt à moyen terme. Le principe est le même que pour le courrier électronique, mais par fax ou par courrier postal. Bien que cette technique soit plus lente, elle possède un gros avantage sur le courrier électronique. En effet, il se fait qu'actuellement le milieu de la sidérurgie n'est pas encore fort présent sur InterNet. En tout cas, pas encore assez pour que les grandes industries y déposent publiquement des informations détaillées, concernant leur fonctionnement interne, et en particulier leurs projets de recherche. On les comprend aisément ... Secret industriel oblige ! Le problème est moindre pour la poste, car de nos jours, on peut joindre pratiquement n'importe quelle personne, du moment qu'elle possède une adresse (postale).

Enfin la dernière manière, est une recherche « aveugle » dans des revues et livres spécialisés en mathématique floue, ... Aveugle en ce sens que l'on ne dispose pas toujours sous la main des bonnes revues, des bons livres. Même si l'on en dispose, il nous faut y trouver des articles se rapprochant le plus possible de ce que l'on cherche, pour une application industrielle (et donc pas uniquement théorique). Le gros avantages des revues

spécialisées est que l'information y est en général « sûre », précise, et abondante, mais que, par conséquent, on y trouve où on y trouve pas ce que l'on cherche. Nous disposons, à la Bibliothèque Universitaire Moretus Plantin de Namur, de la revue « fuzzy sets and systems » depuis le début des années quatre-vingt, et d'un excellent livre d'introduction à la théorie des ensembles flous et ses applications, écrit par H.-J. Zimmerman, un des pionniers. Le résultat des deux premières recherches nous permettra probablement de diminuer notre cécité, afin que nous puissions nous guider plus facilement et plus rapidement, sans donner des coups de machette n'importe où.

Dans les parties qui suivent, nous donnons le résultat de recherches intensives, s'étalant sur une période de trois mois, d'octobre à décembre 1995. Nous essayons en même temps de nous familiariser avec la théorie des ensembles flous. Comme nous l'avons déjà dit, ces recherches sont surtout orientées vers le problème inverse. Il est en effet le seul des trois ne possédant pas encore de solution industrielle implémentée. Il est a fortiori le plus intéressant du point de vue de profilARBED et du nôtre.

4.3 Les résultats des premières recherches

4.3.1 Sur InterNet

Nous avons principalement trouvé deux outils très intéressants sur InterNet : le NewsGroup « fuzzy » et le courrier électronique.

Un « NewsGroup » est une sorte de carnet public, sur lequel tous le monde peut écrire ce qu'il veut. Ce carnet traite en général d'un thème particulier. On en trouve actuellement des milliers dont les thèmes varient du fan club de Madonna, à la recherche sur les réseaux de Petri flous, en passant par les J.O. d'Atlanta. Il est évident que le sérieux d'un NewsGroup dépend de son thème et de ses adhérents. Lorsqu'il est sérieusement utilisé, la coutume veut que l'on y écrive deux types de messages : des questions (signées), et des réponses (signées) à ces questions. C'est un peu comme un dialogue entre personnes où tout le monde peut écouter tout le monde. Le thème de la NewsGroup «fuzzy» est dédié aux mathématiques floues en général. Nous y avons posé trois questions, chacune relative à un des trois problèmes posés. Nous avons reçu pas mal de réponses, dont les plus intéressantes sont des références d'articles ou des adresses de sites.

Le courrier électronique est un outil fort intéressant ... à condition de connaître des adresses électroniques de personnes spécialisées dans les mathématiques floues ou la sidérurgie. En ce qui concerne la sidérurgie, comme nous l'avons déjà dit, c'est apparemment très difficile. Par contre, nous disposons d'une liste très intéressante reprenant les coordonnées (électroniques et postales), ainsi que les spécialités dans les mathématiques floues, de 250 personnes dans le monde entier, dont la plupart sont des chercheurs universitaires. Une version mise à jour de cette liste est disponible en se connectant selon le protocole FTP sur le site « <ftp://mira.dbai.tuwien.ac.at/pub/mlowner/whoiswhoinfuzzy> ». Cette liste constitue un très bon point de départ pour entamer des recherches dans le domaine du flou.

Plus particulièrement pour une résolution du problème de planification de production, basée sur les réseaux de Petri flous, il existe une « mailing list » dont le thème y est consacré.

L'adresse électronique de son coordinateur est « fuzzy ». Une « mailing-list » est une simulation, un prototype de NewsGroup, géré à partir du courrier électronique. On la crée souvent pour voir s'il serait utile de créer le NewsGroup correspondant. Elle possède aussi un thème, est gérée par un coordinateur, et est constituée d'une liste de personnes s'intéressant à ce thème. Lorsqu'une personne de cette liste veut poser (répondre à) une question sur le thème, elle envoie sa question (réponse) au coordinateur. Si cette question (réponse) est jugée sérieuse par ce dernier, il se charge de la communiquer à tous les membres de la liste. Nous recevons actuellement en moyenne deux courriers électroniques par jour concernant la « mailing-list » « fuzzy ».

4.3.2 *Par courrier et fax*

Etant donné le peu de temps que nous avons devant nous, nous avons essayé principalement de joindre H.-P. Lipp et W. Limbeck (problème de planification) pour leur demander s'ils n'avaient pas écrit d'articles en anglais traitant du même sujet que ceux que nous avions en Allemand. Nous avons ainsi reçu une réponse de Monsieur Limbeck. Il stipule que lui-même n'avait pas écrit d'article en anglais sur ce sujet, et que sa connaissance sur les réseaux de Petri flous est faible. Il nous a aussi confirmé que l'entreprise Hüttenwerke Krupp Mannesmann GmbH développe bien une extension floue d'un logiciel permettant d'aider la production de charges d'acier, très similaire au problème de planification de profilARBED. Cette application fut principalement développée par H.-P. Lipp, qui travaille actuellement comme professeur à l'école technique de Schmalkalden, et habite à une nouvelle adresse, qui nous fut d'ailleurs communiquée par W. Limbeck. Par la suite, nous n'avons plus essayé de joindre H.-P. Lipp, étant donné que nous avions, à ce moment, déjà décidé de ne plus nous pencher sur le problème de planification de production.

4.3.3 *Dans les revues et livres spécialisés*

Les articles intéressants concernant les trois problèmes posés sont repris dans la bibliographie. Pour le problème direct, nous disposons d'un livre très intéressant, présentant, et comparant différentes approches de la résolution de PL probabilistes et possibilistes. Ce problème est aussi très fréquemment traité dans la revue « fuzzy sets and systems ».

Notre recherche fut axée principalement sur le problème inverse. Grâce aux deux premières recherches et quelques jours passés dans les bibliothèques, nous avons trouvé assez bien d'articles sur la résolution de régressions linéaires floues ou possibilistes. Ce domaine a principalement été étudié dans les années quatre-vingt. Le pionnier en est apparemment H. Tanaka, un japonais. La plupart des modèles rencontrés considèrent des régressions linéaires floues, où seuls le vecteur des variables de sorties (dans notre cas A_e), et le vecteur des variables indépendantes (dans notre cas T_e) ont leurs composantes dans $F(\mathcal{R})$ l'ensemble des nombres flous. Ils expliquent le fait que l'égalité $Q \otimes T_e = A_e$ ne peut être vérifiée dans le réel en introduisant « du flou » sur les variables de sortie et les variables indépendantes. A priori, ces modèles ne sont pas intéressants pour nous, étant donné qu'ils ne considèrent pas que la matrice des variables d'entrée (dans notre cas Q) puisse aussi être floue et appartenir à $F(\mathcal{R})^{M \times C}$. Rappelons que dans notre cas, seules les variables d'entrée, soit Q , sont physiquement approximatives.

4.4 La sélection d'un des trois problèmes

Un seul de ces modèles considère le cas où Q peut-être floue. Il est publié dans un article de 1989, et écrit par Sakawa et Yano [SY92]. Ces derniers présentent d'ailleurs leur modèle comme une généralisation des modèles introduits et basés sur ceux de Tanaka. Cet article semble donc convenir parfaitement à une résolution du problème inverse. Mais, avant de nous lancer sur « cette si belle piste » à travers la jungle , nous avons tenté de joindre Sakawa et Yano au Japon. Ceci afin de s'assurer que l'on n'y a pas trouvé, entre 1989 et 1996, des erreurs importantes, et/ou, savoir si les auteurs n'en ont pas écrit de plus récents sur le même sujet. Nous y sommes arrivés, en contactant tous les chercheurs japonais, pour lesquels nous disposions d'une adresse électronique ou postale. Peut-être existe-t'il déjà une application industrielle basée sur cet article. Bien que nous soyons quasi certains que Sakawa ait reçu notre message (par l'intermédiaire d'un de ses collègues qui nous l'a confirmé), nous n'avons jamais reçu de réponse...

Après en avoir discuté avec Monsieur Lescrenier et après avoir compris dans les détails l'article, nous nous sommes décidés à préparer armes et bagages pour notre périple vers la solution Sakawa. Elle nous semble mathématiquement correcte et coller parfaitement à nos désirs. C'est donc avec beaucoup d'enthousiasme que nous nous sommes lancés à la découverte de cette piste.

Chapitre 5

**Le modèle Sakawa-Yano comme
solution au problème inverse**

5. Le modèle Sakawa-Yano comme solution au problème inverse.

5.1 Présentation du modèle

Sakawa et Yano [SY92] se proposent, indépendamment de toute sémantique sur les paramètres, de résoudre un problème du type suivant, grâce à des nombres flous. Ils se placent implicitement dans le contexte de la théorie des ensembles flous possibiliste

Déterminer T_e TQ $Q \otimes T_e \cong A_e$

Ce problème, exprimé sous forme matricielle, peut être formulé de la manière suivante :

Déterminer $t(m,e) \forall m \in M$

$$SC \quad \sum_{m \in M} t(m,e) \times q(c,m) \approx a(e,c) \quad \forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}$$

où les correspondances de notation et les hypothèses utilisées dans l'article de Sakawa et Yano sont expliquées ci-dessous.

Remarquons qu'il ne s'agit pas d'expliquer dans les moindres détails le raisonnement de Sakawa-Yano. Nous désirons plutôt faciliter la tâche d'une personne qui voudrait éventuellement reprendre l'article de Sakawa-Yano pour l'appliquer à la résolution du problème inverse de la même manière que nous. Pour les personnes désireuse d'en savoir plus, nous ne pouvons que leur conseiller de lire cet article. Une photocopie est fournie en annexe. Tous les concepts flous abordés dans cet article, sont introduits et expliqués en profondeur dans le chapitre d'introduction à la théorie des ensembles flous. Pour une bonne compréhension de l'article, nous renvoyons le lecteur à ce chapitre.

5.2 Les hypothèses du modèle

5.2.1 Les correspondances de notation

La correspondance entre les notations Sakawa-Yano et nos conventions de modélisation est exprimée dans le tableau suivant :

Sakawa Yano	profilARBED
$j = 1, 2, \dots, n$	$m = 1, 2, \dots, M \Leftrightarrow m \in \mathcal{M}$
$\tilde{a}_0 = 0$ $x_{0i} = 0$	nous désirons résoudre un régression linéaire de type $Q \ t_e$ $\approx a_e$ et non $Q \otimes t_e \oplus b \approx a_e$
$i = 0, 1, \dots, k$	$c = 1, 2, \dots, C \Leftrightarrow c \in \mathcal{C}$
x_{ij}	$q(c, m)$
\tilde{a}_i	$t(m, e)$
y_i	$a(e, c)$

5.2.2 Les hypothèses mathématiques

D'emblée, Sakawa et Yano ont fait une série d'hypothèses, que nous regroupons ci dessous. Ces hypothèses sont purement mathématiques.

- \times la loi de multiplication de deux nombres flous définie par extension de la loi interne de multiplication de deux nombre réels, selon le principe d'extension de Z.D.P.
- \sum une sommation selon la loi d'addition de deux nombres flous définie par extension de la loi interne d'addition de deux nombres réels.
- \approx une des trois relations d'égalité dérivée des indices d'égalités proposés par Dubois et Prade.
- $L : \mathcal{R}^+ \longrightarrow \mathcal{R}$ une fonction repère continue et strictement décroissante sur $[0,1]$ définie de telle manière que $L(0) = 1$ et $L(1) = 0$. Cette fonction nous permet de définir les sous ensembles $F_{L-L}(\mathcal{R})$ et $F_{L-L}(\mathcal{R})^+$ de l'ensemble des nombre flous $F(\mathcal{R})$ (voir introduction à la théorie des ensembles flous).
- $q(c, m) \in F_{L-L}(\mathcal{R})^+$, symétrique, et est noté $(q_c(c, m), q_d(c, m)) \quad \forall m \in \mathcal{M} \quad \forall c \in \mathcal{C}$
- $a(e, c) \in F_{L-L}(\mathcal{R})$, symétrique, et est noté $(a_c(e, c), a_d(e, c^2)) \quad \forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}$
- $t(m, e) \in F_{L-L}(\mathcal{R})$, symétrique, et est noté $(t_c(m, e), t_d(m, e)) \quad \forall m \in \mathcal{M} \quad \forall e \in \mathcal{E}$

5.3 Satisfiabilité des hypothèses Sakawa-Yano

Nous voulons dans cette partie vérifier que les hypothèses proposées par Sakawa-Yano sont acceptables, si l'on affecte aux variables muettes citées ci-dessus, la sémantique que l'on leur a donnée dans nos conventions de modélisation. Ces conventions ne sont en fait rien d'autre que des hypothèses physiques. Autrement dit, nous désirons simplement vérifier que ces hypothèses physiques sont bien couvertes par les hypothèses mathématiques posées par Sakawa-Yano.

Nous pouvons montrer qu'elles le sont, si nous donnons aux centres et aux dispersions de chacun des paramètres la signification suivante :

5.3.1 Les variables d'entrée

$q(c,m)$	= le poids de la quantité du type de mitraille m utilisée pour la production de la charge c
$q_c(c,m)$	= la valeur mesurée (en tonnes) du poids en question
$q_d(c,m)$	= l'erreur (en tonnes) que l'on estime être faite lors de la mesure de ce poids

Toutes les valeurs comprises entre $(q_c(c,m)-q_d(c,m))$ et $(q_c(c,m)+q_d(c,m))$ sont donc considérées comme possibles. La plus possible (degré d'appartenance égal à 1) étant $q_c(c,m)$ et les moins possibles étant toutes les valeurs plus petites que $(q_c(c,m)-q_d(c,m))$ et plus grandes que $(q_c(c,m)+q_d(c,m))$. Si l'on veut signifier que toutes les valeurs comprises entre $(q_c(c,m)-q_d(c,m))$ et $(q_c(c,m)+q_d(c,m))$ sont pratiquement uniformément possibles, il suffit de choisir la fonction repère « L très concave ». Comme un poids est toujours positif, on a toujours pour $(q_c(c,m)-q_d(c,m))$ est positif. Dans ce cas on a bien $q(c,m) \in F_{L-L}(\mathcal{H})^+$.

5.3.2 Les variables de sortie

$a(e,c)$	= le poids de l'élément e dans la charge produite c (a pour analyse)
$a_c(e,c)$	= la valeur mesurée (en tonnes) du poids en question
$a_d(e,c)$	= l'erreur (en tonnes) que l'on estime être faite lors de la mesure

Toutes les valeurs comprises entre $(a_c(c,m)-a_d(c,m))$ et $(a_c(c,m)+a_d(c,m))$ sont donc considérées comme possibles. La plus possible (degré d'appartenance égal à 1) étant $a_c(c,m)$ et les moins possibles étant toutes les valeurs plus petites que $(a_c(c,m)-a_d(c,m))$ et plus grandes que $(a_c(c,m)+a_d(c,m))$. Si l'on veut signifier que toutes les valeurs comprises entre $(q_c(c,m)-q_d(c,m))$ et $(q_c(c,m)+q_d(c,m))$ sont pratiquement uniformément possibles, il suffit de choisir la fonction repère « L très concave ». On a donc bien, au moins, que $a(e,c) \in F_{L-L}(\mathcal{H})$.

5.3.3 Les variables indépendantes

$t(m,e)$	= teneur de la mitraille m en l'élément e trouvée par le modèle
$t_c(m,e)$	= la valeur centrale (la plus possible) de cette teneur
$t_d(m,e)$	= la dispersion de cette teneur.

Toutes les valeurs comprises entre $(t_c(m,e)-t_d(m,e))$ et $(t_c(m,e)+t_d(m,e))$ sont donc considérées comme possibles. La plus possible (degré d'appartenance égal à 1) étant $t_c(m,e)$ et les moins possibles étant toutes les valeurs plus petites que $(t_c(m,e)-t_d(m,e))$ et plus grandes que $(t_c(m,e)+t_d(m,e))$. Le choix de $q(c,m)$ et $a(e,c)$ ayant été fait de telle manière qu'ils appartiennent tous deux au moins à $F_{L-L}(\mathcal{H})$ avec \times et $+$ étant des lois internes de multiplication et d'addition dans $F(\mathcal{H})$, on a bien que $t(c,m)$ appartient à $F_{L-L}(\mathcal{H})$. Si, en plus, on a choisi « L très concave » on peut considérer alors que toutes les valeurs comprises entre $(t_c(m,e)-t_d(m,e))$ et $(t_c(m,e)+t_d(m,e))$ sont pratiquement uniformément possibles.

5.4 Les principales simplifications du modèle Sakawa-Yano

En vertu de la sémantique physique que nous venons d'accorder aux différents paramètres du modèle Sakawa-Yano, nous pouvons faire des hypothèses encore plus fortes que celles proposées par ces auteurs.

On considère que les mesures faites en sortie du circuit de production sont connues au millième près. Autrement dit l'erreur commise sur les analyses est quasi nulle. On peut donc considérer que

$$a_d(e,c) = 0 \text{ et } a_c(e,c) \in \mathcal{R}^+ \Rightarrow a(e,c) \in \mathcal{R}^+ \quad \forall e \in \mathcal{E} \quad \forall c \in \mathcal{C}$$

Il s'en suit donc une simplification drastique des équations du modèle. Nous les citons ici une par une.

5.4.1 Les partitions J1, J2 et J3 sont connues

On sait à l'avance que tous les paramètres du modèle ne peuvent prendre que des valeurs positives. Autrement dit, tous les valeurs possibles appartiennent à \mathcal{R}^+ .

$$J1 \equiv \mathcal{M}$$

$$J2 \equiv \emptyset$$

$$J3 \equiv \emptyset$$

5.4.2 L'expression des α -coupes est plus simple

Il s'agit d'une simple conséquence de ce qui vient d'être dit. L'expression des sommes ne doit plus être découpée selon les ensembles J1, J2 et J3.

$$\forall \alpha \in]0,1] \subseteq \mathcal{R}$$

$$\left(\sum_{m \in \mathcal{M}} t(m,e) \times q(c,m) \right)_{\alpha} = [LB_c^{\alpha}, LU_c^{\alpha}] \quad \forall c \in \mathcal{C}$$

où

$$LB_c^{\alpha} = \sum_{m \in \mathcal{M}} (q_c(c,m) - L^{-1}(\alpha) q_d(c,m)) (t_c(m,e) - L^{-1}(\alpha) t_d(m,e)) \quad \forall c \in \mathcal{C}$$

$$LU_c^{\alpha} = \sum_{m \in \mathcal{M}} (q_c(c,m) + L^{-1}(\alpha) q_d(c,m)) (t_c(m,e) + L^{-1}(\alpha) t_d(m,e)) \quad \forall c \in \mathcal{C}$$

5.4.3 L'expression de la fonction objectif commune

Remarquons que Sakawa et Yano définissent cette fonction par l'intermédiaire de $L^{-1}(0)$, ce qui n'a pas de sens car la fonction $L(\cdot)$ n'est pas inversible. En réalité, il est sous-entendu que Sakawa et Yano restreignent cette fonction à l'intervalle $[0,1]$ et dans ce cas $L^{-1}(0) = 1$. On peut alors reconsidérer l'équation ci-dessus avec α égal à 0. Dans ce cas, l'intervalle $[LB_c^0, LU_c^0]$ n'est autre que le support topologique S_T du nombre flou

$$S_T\left(\sum_{m \in M} t(m,e) \times q(c,m)\right) = [LB_c^0, LU_c^0] \quad \forall c \in \mathcal{C}$$

avec

$$LB_c^0 = \sum_{m \in M} (q_c(c,m) - q_d(c,m)) (t_c(m,e) - t_d(m,e)) \quad \forall c \in \mathcal{C}$$

$$LU_c^0 = \sum_{m \in M} (q_c(c,m) + q_d(c,m)) (t_c(m,e) + t_d(m,e)) \quad \forall c \in \mathcal{C}$$

La fonction objectif commune aux trois problèmes est alors la suivante

$$FO2 = \sum_{c \in \mathcal{C}} (LU_c^0 - LB_c^0)$$

5.4.4 Les deux premiers problèmes sont les mêmes

Intuitivement, on peut dire qu'essayer d'égaliser un nombre flou et un nombre réel revient à essayer d'inclure le nombre réel dans le nombre flou. Mathématiquement on a que

$$\begin{aligned} \text{Pos} \quad & \left(\sum_{m \in M} t(m,e) \times q(c,m) = a(e,c)\right) \geq \alpha \\ \Leftrightarrow \text{Nec} \quad & \left(\sum_{m \in M} t(m,e) \times q(c,m) \supset a(e,c)\right) \geq \alpha \\ \Leftrightarrow & a(e,c) \geq LB_c^\alpha \quad \text{et} \quad LU_c^\alpha \geq a(e,c) \end{aligned}$$

5.4.5 Le problème $P3(\alpha)$ n'admet sémantiquement pas de solution

Intuitivement, on peut dire qu'essayer d'inclure un nombre flou dans un nombre réel est impossible, à condition que le nombre flou ne puisse être réel lui-même. Mathématiquement nous pouvons faire le raisonnement suivant :

L'opérateur \times n'est qu'une généralisation de la multiplication de réels. On a par conséquent l'équivalence suivante, signifiant que la sommation de nombres flous donne un nombre réel

$$\sum_{m \in \mathcal{M}} t(m,e) \times q(c,m) = LB_c^\alpha = LU_c^\alpha \quad \forall c \in \mathcal{C}$$

$$\Leftrightarrow t(m,e), q(c,m) \in \mathfrak{R} \text{ et } t(m,e) \text{ est une constante en tant que nombre réel } (\forall m \in \mathcal{M} \forall c \in \mathcal{C})$$

Or, si nous essayons de résoudre le problème inverse, c'est justement parce que cette proposition n'a pas de sens. Dès lors, on peut considérer que dans les cas qui nous concernent, si le modèle est adéquat, il conclura toujours que $t(m,e)$ est un nombre purement flou. Ce flou est censé exprimer la variation de la teneur réelle. On aura donc toujours

$$LB_c^\alpha < LU_c^\alpha \quad \forall \alpha \in]0,1] \subseteq \mathfrak{R}$$

Comme, dans notre cas, le troisième problème est tel que son convexe des contraintes est défini de la façon suivante :

$$LB_c^{1-\alpha} \geq a(e,c) \geq LU_c^{1-\alpha}$$

$$t_c(m,e) - L^{-1}(\alpha) t_\bullet(m,e) \geq 0$$

avec α dans $]0,1]$

Le convexe des contraintes du troisième problème est donc vide et le PL n'admet pas de solution.

5.5 Application de la solution Sakawa-Yano

Maintenant que nous avons donné une sémantique physique aux paramètres du modèle Sakawa-Yano, toutes les expressions physiques énoncées dans leur article peuvent être physiquement interprétées. Afin de mieux comprendre l'idée des auteurs, nous donnons cette interprétation physique pour deux grandeurs clés du modèle Sakawa-yano :

$$a'(e,c) \quad \text{nombre flou égal à } \sum_{m \in \mathcal{M}} t(m,e) \otimes q(c,m)$$

analyse (approximative) virtuelle de la charge c en l'élément e, c.-à-d. le poids en l'élément e qu'il y aurait dans la charge c si la teneur (approximative) en l'élément e de chaque type de mitraille était $t(m,e)$

$$LU_c^0 - LB_c^0 \quad \text{une mesure de flou de la grandeur } a'(e,c)$$

la longueur de l'intervalle des valeurs possibles, balayé par l'analyse virtuelle de la charge c en l'élément e ou encore l'erreur possible sur l'analyse virtuelle de la charge c en l'élément e.

$$\text{Pos}(a'(e,c)) = a(e,c) \quad \text{un indice d'égalité entre } a'(e,c) \text{ et } a(e,c)$$

Indice de proximité entre l'analyse virtuelle (approximative) de la charge c en l'élément e et l'analyse réelle (précise) de la charge c en l'élément e

En tenant compte des différentes simplifications qui sont mentionnées dans la partie précédente, la solution Sakawa-Yano appliquée au problème inverse est alors basée sur la résolution des E problèmes d'analyse multicritère suivants, chacun étant propre à un élément e :

$$\forall e \in \mathcal{E}$$

$$\text{MAX FO1} = \alpha \quad \text{et} \quad \text{MIN FO2} = \sum_{c \in \mathcal{C}} (LU_c^0 - LB_c^0)$$

$$\begin{aligned} \text{SC} \quad & \text{Pos} \quad (a'(e,c) = a(e,c)) \geq \alpha & \forall c \in \mathcal{C} \\ & t_c(m,e) \geq 0 & \forall m \in \mathcal{M} \\ & t_c(m,e) - L^{-1}(\alpha) t_d(m,e) \geq 0 & \forall m \in \mathcal{M} \\ & 0 < \alpha \leq 1 \end{aligned}$$

« Pour chaque élément e, déterminer la teneur approximative de chaque type de mitraille en cet élément e de telle manière que, pour chaque charge produite, l'indice d'égalité entre l'analyse réelle de cette charge en l'élément e et l'analyse virtuelle de cette charge en l'élément e, soit plus grand qu'une valeur la plus élevée possible. Par définition des indices d'égalité, il existe une infinité de valeurs floues vérifiant ces contraintes. On choisit celles qui minimisent la somme, sur l'ensemble des charges, des erreurs possibles sur l'analyse virtuelle. »

Comme nous l'avons montré, chacune de ces expressions est linéaire en t(m,e). Ces E problèmes d'analyse multicritères, notés PLM₁(e), sont en fait E PL multicritères (deux fonctions objectifs). Le concept mathématique de solution optimale d'un PL monocritère (une seule fonction objectifs) n'a évidemment plus de sens. Par exemple, si t_e^s (resp α^s = 0) est la solution optimale du PL obtenu en retirant FO1 (resp FO2) de PLM₁(e), alors le couple(t_e^s, α^s) n'appartient pas au convexe des contraintes de PLM(e).

Pour palier à cela, Sakawa et Yano utilisent la notion de solution optimale de Pareto d'un problème multicritère.

Par définition, le couple (t_e^p, α^p) est une **solution optimale de Pareto** de PLM(e) ssi il n'existe pas une autre couple (t_e, α) tel que FO1(α) = α ≥ α^p = FO1(α^p) et FO2(t_e) ≤ FO2(t_e^p) avec une inégalité stricte pour au moins une des deux propositions.

Il y a mathématiquement toujours une infinité de solutions optimales de Pareto. Or il est fort probable qu'une seule ou une partie seulement des solutions nous intéressent. Il s'agit de celle ou celles qui sont représentatives de l'évolution de la teneur de chaque type de mitraille sur l'ensemble des charges produites.

Pour tenter de générer les solutions optimales de Pareto qui nous intéressent, Sakawa et Yano disent qu'il est « pratique » de suivre la méthode suivante. Supposons que nous fixions la valeur de α à α₀ (choisi dans l'intervalle]0,1]) pour chaque PLM₁(e), et que nous le

remplacements par un PL dépendant de deux paramètres e et α_0 . Chacun d'eux deviendrait alors un PL classique noté $PL_1(\alpha_0, e)$ suivant :

$$\begin{aligned} \text{MIN } \mathbf{FO2} &= \sum_{c \in \mathcal{C}} (LU_c^0 - LB_c^0) \\ \text{SC } \text{Pos } (\mathbf{a}'(\mathbf{e}, \mathbf{c}) = \mathbf{a}(\mathbf{e}, \mathbf{c})) &\geq \varepsilon \quad \forall c \in \mathcal{C} \\ \varepsilon &\geq \alpha_0 \\ \mathbf{t}_c(\mathbf{m}, \mathbf{e}) &\geq 0 \quad \forall m \in \mathcal{M} \\ \mathbf{t}_c(\mathbf{m}, \mathbf{e}) - L^{-1}(\varepsilon) \mathbf{t}_d(\mathbf{m}, \mathbf{e}) &\geq 0 \quad \forall m \in \mathcal{M} \end{aligned}$$

La relation entre $PL_1(\alpha_0, e)$ et $PLM_1(e)$ est donnée par Sakawa et Yano.

$$\forall e \in \mathcal{E}$$

\mathbf{t}_e^s est une solution optimale unique de $PL_1(\alpha_0, e) \Leftrightarrow (\mathbf{t}_e^s, \alpha_0)$ est une solution optimale de Pareto de $PLM_1(e)$.

On remarque en quoi ce raisonnement est pratique. Etant donné l'implication « \Rightarrow » si l'on résout $PL_1(\alpha_0, e)$ et que l'on trouve une solution unique, alors on a trouvé un optimal de Pareto. De même, en se servant de l'implication « \Leftarrow », on en déduit que si l'on résout $PL_1(\alpha_0, e)$ et que l'on a pas une solution unique (c.-à.-d. pas de solution ou une infinité de solutions), alors on n'a pas d'optimal de Pareto en appliquant la construction. En balayant différentes valeurs de α_0 sur $]0, 1]$ pour $PL_1(\alpha_0, e)$, on balaie une partie (pas toutes) des solutions optimales de Pareto $PLM_1(e)$. Il existe plusieurs méthodes qui tentent de faire de même, mais actuellement aucune n'est apparemment sûre. Nous remarquerons plus tard que nous obtenons une solution unique dans la plupart des cas.

Suite à cette relation, afin que l'utilisateur, puisse se focaliser sur la ou les solutions optimales de Pareto qui l'intéressent, Sakawa-Yano proposent un algorithme interactif qui constitue notre solution au problème inverse :

$$\forall e \in \mathcal{E}$$

1	Demander à l'utilisateur une valeur de α_0 dans l'intervalle $]0, 1]$. Passer à l'étape 2.
2	Résoudre $PL_1(\alpha_0, e)$. Soit \mathbf{t}_e^s la solution trouvée et $\mathbf{FO2}^s$ la valeur de la fonction objectif : <ul style="list-style-type: none"> • \mathbf{t}_e^s n'est pas une solution unique. Annoncé « tentative infructueuse » et retourner à l'étape 1. • \mathbf{t}_e^s est une solution unique. Dans ce cas passer à l'étape 3.
3	<ul style="list-style-type: none"> • l'utilisateur trouve que $\mathbf{FO2}^s$ n'est pas satisfaisante pour α_0. Retourner à l'étape 1. • l'utilisateur trouve que $\mathbf{FO2}^s$ est satisfaisante pour α_0. Les teneurs de chaque type de mitraille m en l'élément e sont données par les M nombres $\mathbf{t}(\mathbf{m}, \mathbf{e})$ appartenant à $F_{L-L}(\mathcal{M})$ où $\mathbf{t}(\mathbf{m}, \mathbf{e}) = [\mathbf{t}_e^s]_m$.

Nous donnons ci dessous l'expression détaillée du $PL_1(\alpha_0, e)$

$$\text{MIN} \quad \text{FO2} = \sum_{m \in \mathcal{M}} \left[\left(2 \sum_{c \in \mathcal{C}} q_d(c, m) \right) t_c(m, e) + \left(2 \sum_{c \in \mathcal{C}} q_c(c, m) \right) t_d(m, e) \right]$$

$$\text{SC} \quad \sum_{\substack{\forall c \in \mathcal{C} \\ m \in \mathcal{M}}} (q_c(c, m) - L^{-1}(\alpha_0) q_d(c, m)) (t_c(m, e) - L^{-1}(\alpha_0) t_d(m, e)) \leq a(e, c)$$

$$\sum_{\substack{\forall c \in \mathcal{C} \\ m \in \mathcal{M}}} (q_c(c, m) + L^{-1}(\alpha_0) q_d(c, m)) (t_c(m, e) + L^{-1}(\alpha_0) t_d(m, e)) \geq a(e, c)$$

$$\begin{aligned} t_c(m, e) - L^{-1}(\alpha_0) t_d(m, e) &\geq 0 & \forall m \in \mathcal{M} \\ t_c(m, e) &\geq 0 & \forall m \in \mathcal{M} \end{aligned}$$

Sakawa fournit aussi un indice d'orientation pour l'utilisateur afin que celui-ci puisse savoir s'il doit augmenter ou diminuer la valeur de α_0 quand il passe de l'étape 3 à l'étape 1.

5.6 Les premières critiques de la solution Sakawa-Yano

Bien que la plupart des critiques de la solution Sakawa-Yano se feront dans la partie sur les tests de la solution, nous pouvons d'emblée émettre quelques critiques.

Celle qui vient en premier à l'esprit, est que les auteurs prétendent trouver la solution générale au problème de régression linéaire floue, indépendamment de toute sémantique associée aux variables. C'est exactement comme si quelqu'un disait qu'il a trouvé la meilleure vitesse à adopter pour une voiture. Dans l'absolu, ça n'a pas de sens. En fait, la solution Sakawa-Yano est celle qui correspond à un Optimal de Pareto, exactement comme un régression au moindre carré est celle qui correspond à une norme 2. Il se fait que la norme deux ne convient pas pour le problème inverse. L'optimal de Pareto conviendra peut-être.

De plus, Sakawa et Yano ne discutent pas plus à fond le fait que la résolution de $PL_1(\alpha_0, e)$ peut donner une solution multiple. En y réfléchissant un peu, on remarque que dans ce cas, on peut faire certaines déductions supplémentaires. Peut-être même y a-t-il une erreur, ce qui expliquerait pourquoi Sakawa n'a jamais répondu aux appels que nous lui faisions.

Enfin, nous avons constaté que Sakawa et Yano veulent absolument trouver la valeur de α la plus grande possible correspondant à la solution qui nous intéresse le plus. Etant donné qu'ils ne donnent aucune sémantique à leurs variables, cela paraît logique. Cependant, on peut imaginer que l'affectation d'une sémantique particulière aux variables peut aller tout-à-fait à l'encontre d'une augmentation de α . Nous en aurons une preuve éclatante pour notre application. En effet, nous constaterons que la solution qui nous intéresse est justement celle qui correspond à une valeur de α égale à 0 !

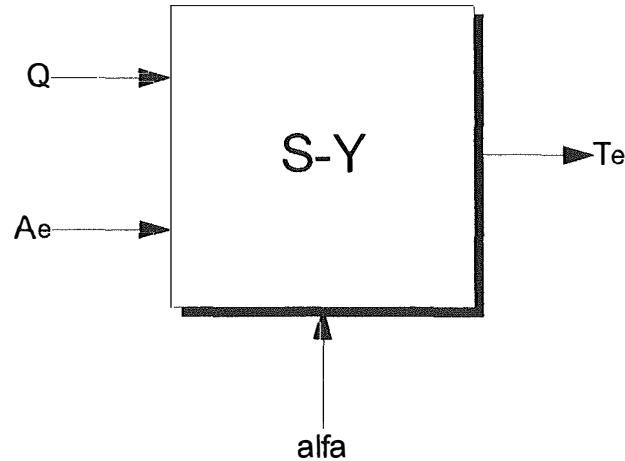
Chapitre 6

Tests de la solution Sakawa-Yano

6. Tests de la solution Sakawa-Yano

6.1 Comment tester le modèle ?

Pour résumer tout ce qui a été dit jusqu'à présent, nous disposons d'un modèle que l'on peut schématiser de la manière suivante :



En entrée du modèle, nous disposons de trois données : la matrice Q , dite des variables d'entrée, le vecteur A_e , dit des analyses, et la valeur α , dite indice de régression, entre 0 et 1. Rappelons que

Les conventions

$q(c,m) \in F_{L-L}(\mathcal{H})^+$	et est noté	$(q_c(c,m), q_d(c,m))$	$\forall m \in \mathcal{M} \forall c \in \mathcal{C}$
$a(e,c) \in F_{L-L}(\mathcal{H})^+$	et est noté	$(a_c(e,c), a_d(e,c))$	$\forall e \in \mathcal{E} \forall c \in \mathcal{C}$
$[Q]_{c,m} = q(c,m)$	$\forall m \in \mathcal{M} \forall c \in \mathcal{C}$		
$[A_e]_c = a(e,c)$	$\forall e \in \mathcal{E} \forall c \in \mathcal{C}$		

Les variables d'entrée

$q(c,m)$ = le poids de la quantité du type de mitraille m utilisée pour la production de la charge c

$q_c(c,m)$ = la valeur mesurée (en tonnes) du poids en question

$q_d(c,m)$ = l'erreur (en tonnes) que l'on estime être faite lors de la mesure de ce poids

Les variables de sortie

$a(e,c)$ = le poids de l'élément e dans la charge produite c (a pour analyse)

$a_c(e,c)$ = la valeur mesurée (en tonnes) du poids en question

$a_d(e,c)$ = 0 (l'erreur (en tonnes) que l'on estime être faite sur de cette mesure)

L'indice de régression

α déterminé par l'utilisateur, compris entre 0 et 1

Le modèle nous renvoie en sortie, un vecteur T_e , dit des variable indépendantes.

Les variables indépendantes

$t(m,e) \in F_{L-L}(\mathcal{H})$ et est noté $(t_c(m,e), t_d(e,c)) \quad \forall e \in \mathcal{E} \quad \forall m \in \mathcal{M}$

$t(m,e)$ = teneur de la mitraille m en l'élément e trouvée par le modèle

$t_c(m,e)$ = la valeur centrale (la plus possible) de cette teneur

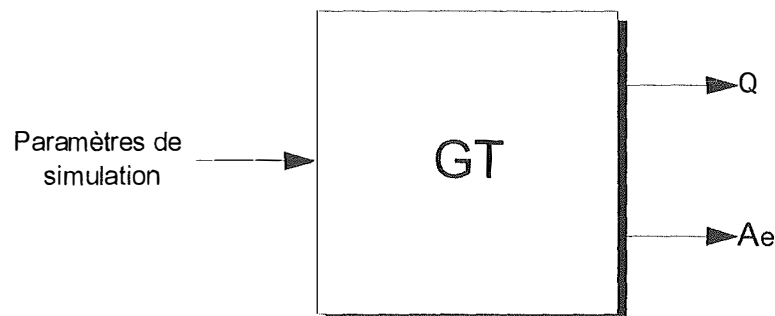
$t_d(m,e)$ = la dispersion de cette teneur.

$[T_e]_m = T(m,e) \quad \forall e \in \mathcal{E} \quad \forall m \in \mathcal{M}$

Mathématiquement, en fonction de ce qui à été dit dans le chapitre précédent, on peut symboliser T_e comme étant le vecteur qui vérifie l'égalité $Q \otimes T_e \cong A_e$ au sens d'un indice d'égalité entre nombres flous, inventé par Dubois et Prade, lié à l'indice de régression. Physiquement, on espère que tout intervalle $[t_c(m,e)-t_d(m,e), t_c(m,e)+t_d(m,e)]$, trouvé par le modèle, est représentatif de l'intervalle de variation sur l'ensemble des charges, de la teneur en l'élément e du type de mitraille m . Le but de cette partie est de vérifier si le modèle se comporte de cette manière.

Cependant, afin de réaliser cet objectif, nous ne pouvons pas travailler directement à partir de données de production. En effet, supposons que l'on le fasse et que le modèle nous dise que l'intervalle de variation de la teneur du type de mitraille 6 en fer sur l'ensemble des charges, est $[0.90, 0.96]$. Nous n'avons aucun moyen de savoir si cette assertion est vraie. Il nous faudrait disposer de données virtuelles de production pour lesquelles on connaît exactement, à l'avance, comment les teneurs de chaque type de mitraille en chaque élément varient sur l'ensemble des charges. Ainsi, si on sait à l'avance que, sur l'ensemble des charges produites, la teneur du type de mitraille 6 en fer a varié dans l'intervalle $[0.90, 0.96]$, alors on peut en déduire que le modèle trouve bien l'intervalle que l'on voulait.

Pour créer ces données virtuelles de production, nous avons construit un générateur de test. Ce générateur simule la production d'un certain nombre de charges. Pour un élément e donné, il génère une matrice Q (dont les composantes sont de l'ordre de grandeur des valeurs industrielles), calcule le vecteur A_e qui lui correspondrait, pour un comportement connu de variations des teneurs de chaque type de mitraille, et modifie la matrice Q pour tenir compte des erreurs de mesures évoquées précédemment. On peut le schématiser de la manière suivante :



En entrée, il dispose d'un certain nombre de paramètres de simulations de production. Ceux-ci décrivent principalement le comportement de variation des teneurs sur l'ensemble des charges, de chaque type de mitraille, en chaque élément, et les erreurs commises sur la mesure du poids des mitrailles. Nous expliquerons plus en détail ce générateur dans la partie

suivante. En sortie, ce générateur renvoie des données virtuelles de production, pour lesquelles on connaît à l'avance les intervalles de variations des teneurs.

6.2 Le générateur de test

6.2.1 Les paramètres de simulation

Nous détaillons dans cette partie les différents paramètres de simulation. Nous pouvons en distinguer trois catégories. Les paramètres globaux, qui décrivent en quelque sorte la « quantité » des données virtuelles de production à générer. Il y a ensuite les paramètres servant à décrire la façon dont la teneur des différents types de mitraille varient sur l'ensemble des charges, et enfin, un paramètre décrivant l'erreur faite sur la mesure des poids des mitrailles.

6.2.1.1 Les paramètres globaux

On considère qu'il n'y a qu'un élément chimique, le fer, pour lequel on utilise l'indice e. Ensuite nous avons

M =	nombre de types de mitraille
C =	nombre de charges à produire à partir de ces types de mitraille
NMC =	nombre maximum de mitrailles par charge, les M mitrailles ne sont pas toutes utilisées pour la production d'une charge.

6.2.1.2 Les paramètres du comportement de variation des teneurs

Pour chacune des M types de mitraille, on caractérise son comportement de variation, sur l'ensemble des charges dans lesquelles on va s'en servir. Chaque type de mitraille est décrit par les paramètres suivants :

m =	indice du type de mitraille
IV(m) =	$[IV(m)_{\min}, IV(m)_{\max}] \subseteq \mathcal{R}$ = intervalle de variation du type de mitraille m en fer sur l'ensemble des charges dans lesquelles on s'en sert.
MV(m) \in {B,S} =	modes de variation de cette teneur dans l'intervalle IV(m).
IV(m) _{min} = IV(m) _{max} \Leftrightarrow	MV(m) ne prend pas de valeur

Nous utilisons principalement deux modes de variation. Il y a le modes binaire, noté « B », signifiant que la teneur du type de mitraille m varie dans l'intervalle IV(m), en prenant une fois sur deux la valeur IV(m)_{min}, et une fois sur deux la valeur IV(m)_{max}. Il y a le mode sinusoïdal, noté « S », signifiant que la teneur du type de mitraille m varie sinusoïdalement sur l'ensemble des charges produites,

6.2.1.3 Les paramètres du comportement des erreurs commises sur les mesures

Si l'on se place du point de vue même de la production, a priori, on ne dispose pas encore d'une description fine de la façon dont l'erreur de mesure des poids est commise. Par

exemple, on ne peut pas savoir si l'erreur est plus grande lorsqu'il s'agit d'une grosse pesée, ou encore lorsqu'il est tard et que les ouvriers de production sont fatigués. Nous décidons donc de ne pas caractériser cette erreur, mitraille par mitraille comme on vient de la faire pour les paramètres de variation, mais plutôt de manière globale pour toutes les pesées. Nous n'avons donc qu'un paramètre d'erreur, noté Err, compris entre 0 et 100 pour relater un pourcentage d'erreur.

6.2.1.4 La fonction $L(.)$ et la dispersion sur les nombres flous

Arbitrairement, nous choisissons la fonction $L(.)$ de la manière suivante

$$L : \mathbb{R}^+ \longrightarrow [0,1] \subseteq \mathbb{R} \\ x \qquad \qquad L(x) = \text{Max}(1-x, 0)$$

Nous manipulerons ainsi des nombres flous de type L-L symétriques et triangulaires. Soit $q_c(c,m)$ le centre d'un de ces nombres, nous choisissons une dispersion $q_d(c,m)$ égale à $0.1 \times q_c(c,m)$. Nous considérons donc que les poids mesurés sont connus à 10% près.

Remarquons une particularité pratique de cette fonction. Si on la restreint à l'intervalle $[0,1]$ elle devient inversible. Soit $L^{-1}(.)$ son inverse alors

$$L(x) = L^{-1}(x) = 1-x \quad \forall x \in [0,1] \subseteq \mathbb{R}$$

6.2.2 La politique de génération virtuelle des charges

Nous voulons ici expliquer brièvement, sous la forme d'un pseudo algorithme, la manière dont la matrice Q et le vecteur A_e sont générés. Pour chacune des C charges, indicées par c , nous procédons de la manière suivante.

1. Choisir aléatoirement NMC types de mitraille m parmi les M disponibles, selon une distribution uniforme de probabilité sur \mathcal{M} .
2. Pour chacun d'eux
 - déterminer son poids réel (puisé dans la cuve) $q_c(c,m)$ dans l'intervalle $[20, 100]$ tonnes, afin que la somme totale des poids dans cette charge vaille 100 tonnes.
 - déterminer sa teneur réelle en l'élément e ($tr(m,e,c)$) dans cette charge, en fonction de ses bornes de variation et de son mode de variation.
3. Calculer la quantité $a(e,c)$, en l'élément e contenue dans la charge c , en fonction des NMC types de mitraille m choisis pour cette charge, de leur quantité $q_c(c,m)$, et de leur teneur dans cette charge $tr(m,e,c)$. Il s'agit d'une somme des $q_c(c,m)$, sur les NMC types de mitraille choisis m , pondérée par les $tr(m,e,c)$
4. Modifier les poids $q_c(c,m)$, selon une distribution uniforme de probabilité sur l'intervalle $[q_c(c,m)(1-Err/100), q_c(c,m)(1+Err/100)]$. Il devient ainsi le poids mesuré.
5. A partir des nouvelles valeurs de $q_c(c,m)$, déterminer leur flou correspondant ($q_d(c,m)$) selon une distribution uniforme dans l'intervalle $[0.9 \times q_c(c,m), 1.1 \times q_c(c,m)]$

Nous tenons à souligner la façon dont sont générés les flous $q_d(c,m)$ sur les poids mesurés $q_c(c,m)$. En pratiquant de cette manière, nous pouvons prouver que l'intervalle $[q_c(c,m)-q_d(c,m), q_c(c,m)+q_d(c,m)]$ contient le poids réellement utilisé du type de mitraille m dans la charge c pour toute valeur de Err comprise entre 0 et 9%. Rappelons que dans les cas réels de production, on ne connaît pas le poids réellement utilisé de la mitraille m dans la charge c . Pour que cette assertion soit vraie, si cette méthode est utilisée en production, il faudra donc déterminer un flou assez grand sur le poids mesuré, de manière à être sûr que le poids utilisé soit contenu dans l'intervalle correspondant. Nous verrons lors des différents tests ce qu'il se passe si ce flou est trop petit ou trop grand.

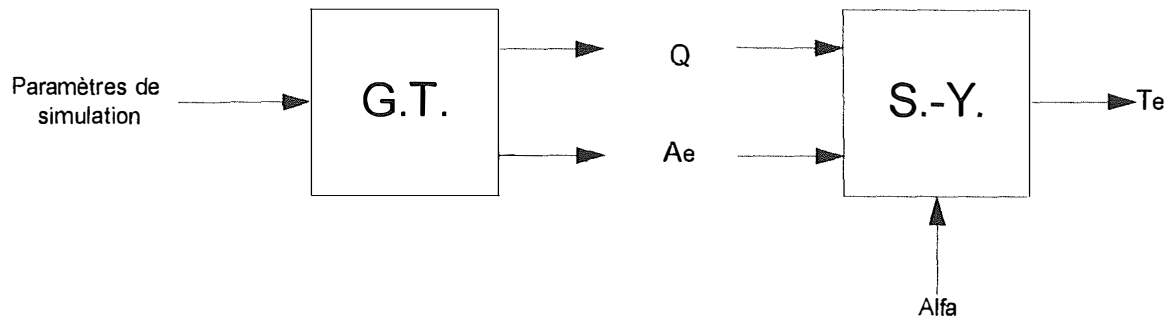
Si l'on supprime la dernière étape de l'algorithme les poids réellement utilisés et les poids mesurés des différents types de mitraille sont représentés par des nombres réels et ne sont pas identiques.. Si l'on supprime les deux dernières étapes de l'algorithme les poids réellement utilisés et les poids mesurés des différents types de mitraille sont représentés par des nombres réels et sont identiques.

6.3 La présentation des résultats

6.3.1 Notre méthodologie de test

Nous avons réalisé une série de tests. Chaque test se réalise de la façon suivante. Nous définissons des paramètres de simulation de production, réalistes par rapport au fer, que nous plaçons en entrée du générateur de test. Ce dernier nous renvoie des données de production Q et A_e , (e étant en l'occurrence l'indice du fer), auxquelles nous appliquons la méthode Sakawa-Yano, pour certaines valeurs de α .

Schématiquement



Nous présentons les résultat de chaque test sous forme de tableau

Selon le test que l'on veut décrire, nous utiliserons soit le tableau 1 pour représenter les résultats, soit le tableau 2. Pour un test donné, le contenu de l'un est toujours déductible de l'autre. Ce sont deux manières différentes de représenter les résultats d'un test. L'utilisation de l'un ou de l'autre sera motivée par ce que l'on désire mettre en évidence dans les résultats d'un test.

6.3.1.1 Le tableau 1

Soit $IV(m)$ l'intervalle de variation que nous avons introduit dans la partie la description des paramètre d'entrée du générateur de test. Soit C son centre et R son rayon.

Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano							
C = ..., NMC = ..., Err = ...%				$\alpha = \dots$		$\alpha = \dots$		$\alpha = \dots$		$\alpha = \dots$	
Etapes 1 à ... de l'algorithme				$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$
m	C	R	MV								
1											
m											
M											

6.3.1.2 Le tableau 2

Appelons $IT(m,e)$ l'intervalle $[t_c(m,e)-t_d(m,e), t_c(m,e)+t_d(m,e)]$. Pour une question évidente de place, nous ne donnerons $IT(m,e)$ que pour 4 valeurs différentes de α au lieu de 10. Nous verrons que cela suffit pour la compréhension du comportement du modèle.

Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C = ..., NMC = ..., Err = ...% Etapes 1 à ... de l'algorithme						
m	IV(m)	MV(m)	IT(m,e) $\alpha = \dots$	IT(m,e) $\alpha = \dots$	IT(m,e) $\alpha = \dots$	IT(m,e) $\alpha = \dots$
1						
m						
M						

Nous avons réalisé grosso modo quatre grands types de tests, regroupant au total une centaine de tests proprement dit du modèle Sakawa-Yano. Nous pourrions donc vous montrer une centaines de tableaux de résultats, ce que nous ne ferons évidemment pas. Nous nous limiterons à quelques tests clés, que nous considérons représentatifs du comportement du modèle Sakawa-Yano.

Les quatre types de test sont par ordre croissant de « complexité ». Les derniers tests sont ceux qui se rapprochent le plus des situations réelles. Les autres sont là simplement pour mieux comprendre le modèle avant de passer aux tests sérieux. Nous avons vu que l'égalité $Q \otimes T_e = A_e$, où les paramètres sont des nombres réels, ne peut être respectée pour deux raisons. D'une part la teneur de chaque type de mitraille en chaque élément n'est pas constante en tant que réel, et d'autre part, les valeurs des composantes de la matrice Q sont erronées. Le modèle Sakawa-Yano, nous renvoie une solution floue au problème, censée « expliquer » au mieux ces deux écarts à l'idéalité. Appelons le premier type d'écart à l'idéalité, **l'erreur de type 1**, et le second **l'erreur de type 2**. Notre générateur de test est conçu pour pouvoir décider à notre guise de générer des données virtuelles de production contenant ou ne contenant pas ces erreurs. Par exemple, nous pouvons générer des données virtuelles de production ne « contenant » pas d'erreurs de type 1 en initialisant chaque IV(m) à une seule valeur, la teneur constante du type de mitraille m. De même, nous pouvons supprimer les erreurs de type 2 en supprimant les deux dernières étapes de l'algorithme. Il y a donc quatre cas de figure différents correspondant aux quatre séries de tests réalisés.

	Données virtuelles de production générées avec	
	erreurs de type 1	erreurs de type 2
Test de type 1	non	non
Test de type 2	oui	non
Test de type 3	non	oui
Test de type 4	oui	oui

Voyons maintenant les résultats proprement dits de ces quatre types de tests.

6.3.2 Les tests de type 1

Nous ne montrons qu'un seul test, car il est très révélateur du comportement observé. Nous nous trouvons ici dans un cas idéal de production. Par construction des données virtuelles de production, dans ce type de test, l'égalité $Q \otimes T_e = A_e$ peut être respectée dans les réels. Le modèle Sakawa-Yano trouve bien la solution qui vérifie cette égalité dans les réels, quelle que soit la valeur de α , même égale à 1.

Test 1.1						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C= 50, NMC = 7, Err= 0% Etapas 1 à 3 de l'algorithme						
m	IV(m)	MV(m)	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = 0.1$	IT(m,e) $\alpha = 0.9$	IT(m,e) $\alpha = 1.0$
1	85.1		85.1	85.1	85.1	85.1
2	85.2		85.2	85.2	85.2	85.2
3	85.3		85.3	85.3	85.3	85.3
4	85.4		85.4	85.4	85.4	85.4
5	85.5		85.5	85.5	85.5	85.5
6	85.6		85.6	85.6	85.6	85.6
7	85.7		85.7	85.7	85.7	85.7
8	85.8		85.8	85.8	85.8	85.8
9	85.9		85.9	85.9	85.9	85.9
10	86.0		86.0	86.0	86.0	86.0

Lorsqu'un intervalle ne contient qu'une valeur V, nous ne le notons pas [V-V] mais simplement V. Nous notons aussi les teneurs en « % » et plus avec un réel compris entre 0 et 1.

6.3.3 Les tests de type 2

Dans ce type de test, nous constatons que la plus représentative toujours celle trouvée pour une valeur de $\alpha = 0$. Cependant, il semble que les solutions trouvées pour différentes valeurs de α soient liées. Une façon de le remarquer est de constater qu'une augmentation de la valeur de α ne modifie pas le centre de IT(m,e) mais augmente le flou linéairement. Nous prouvons ce phénomène dans la partie sur la compréhension des résultats. Nous pouvons donc nous contenter de ne présenter les résultats que pour une seule valeur de α , quelconque dans l'intervalle [0,1]. Nous choisissons $\alpha = 0$ car ce cas est le plus représentatif de la variation des teneurs. Nous pourrions tout aussi bien choisir $\alpha = 0.999$ et vous laisser déduire la solution que l'on aurait obtenue pour $\alpha = 0$ mais ce serait moins parlant. Le problème n'admet jamais de solution pour $\alpha = 1$. Nous montrerons en effet qu'il n'existe une solution dans ce cas que si l'égalité $Q \otimes T_e = A_e$ peut être respectée dans les réels, ce qui n'est pas possible.

Test 2.1											
Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano							
C= 50,NMC = 7,Err =0% Etapas 1 à 3 de l'algorithme				$\alpha = 0$		$\alpha = 0.1$		$\alpha = 0.2$		$\alpha = 0.9$	
m	C(m)	R(m)	MV(m)	$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$	$t_c(m,e)$	$t_d(m,e)$
1	92.550	7.450	---	92.550	7.450	92.550	8.278	92.550	9.313	92.550	74.500
2	85.2	0		85.2	0	85.2	0	85.2	0	85.2	0
3	85.3	0		85.3	0	85.3	0	85.3	0	85.3	0
4	85.4	0		85.4	0	85.4	0	85.4	0	85.4	0
5	85.5	0		85.5	0	85.5	0	85.5	0	85.5	0
6	85.6	0		85.6	0	85.6	0	85.6	0	85.6	0
7	85.7	0		85.7	0	85.7	0	85.7	0	85.7	0
8	85.8	0		85.8	0	85.8	0	85.8	0	85.8	0
9	85.9	0		85.9	0	85.9	0	85.9	0	85.9	0
10	60.0	0		60.0	0	60.0	0	60.0	0	60.0	0

Soit $(t_c^1(m,e), t_d^1(m,e))$, $\forall m \in \mathcal{M}$, la solution trouvée pour α_1

Soit $(t_c^2(m,e), t_d^2(m,e))$, $\forall m \in \mathcal{M}$, la solution trouvée pour $\alpha_2 \neq \alpha_1$

Ces deux solutions sont liées de la manière suivante

$$t_c^1(m,e) = t_c^2(m,e) \quad \forall m \in \mathcal{M}$$

$$t_d^1(m,e) = (L^{-1}(\alpha_1)/L^{-1}(\alpha_2)) t_d^2(m,e) \quad \forall m \in \mathcal{M}$$

Lorsque l'on connaît la solution pour une valeur de α , on peut en déduire la solution pour n'importe quelle autre valeur de α . Comparons par exemple le flou sur la teneur en fer (élément e) du premier type de mitraille ($t_d(1,e)$) pour les solutions obtenues avec les quatre valeurs de α du test 2.1.

$$\begin{aligned} 7.450 &= (L^{-1}(0.1) / L^{-1}(0)) \times 8.278... = (0.9 / 1) \times 8.278... \\ &= (L^{-1}(0.2) / L^{-1}(0)) \times 9.313... = (0.8 / 1) \times 9.313... \\ &= (L^{-1}(0.9) / L^{-1}(0)) \times 74.5 = (0.1 / 1) \times 74.500 \end{aligned}$$

Supposons que l'on se définisse une distance entre la solution trouvée par le modèle Sakawa-Yano pour $\alpha = 0$, et celle attendue, déterminée à l'avance dans les paramètres de simulation de production. Cette distance aurait la propriété d'être nulle lorsque les deux solutions sont identiques. Empiriquement, on constate que, quel que soit le nombre de types de mitraille, il existe un nombre de charges C, au delà duquel cette distance vaut toujours 0. Autrement dit, si l'on produit un assez grand nombre de charges, il est certain que la solution trouvée par le modèle Sakawa-Yano est l'image exacte du comportement de variation des teneurs de chaque type de mitraille en l'élément e.

Remarquons que cette limite inférieure est d'autant plus grande que le nombre de mitrailles variables augmente, mais aussi que l'on évolue d'un mode binaire de variation à un mode sinusoïdal. Par exemple, si l'on refait le même test que le test 2.1 en mode sinusoïdal de variation il faut produire au moins 150 charges pour cette distance vaille 0. Il n'en faut que 50 en modes binaire(test 2.1)! Cette distance est d'autant plus proche de 0 que le nombre de fois que les bornes de variation des teneurs sont physiquement atteintes est grand. Le mode de variation des teneurs a donc une forte importance sur la «vitesse de

convergence » de la solution. Chaque type de mitraille est en moyenne utilisé 10 fois sur 50 charges. Dans le mode de variation binaire les bornes de variations des teneurs sont atteintes chaque fois qu'elles sont utilisées. Dans le mode sinusoïdal elle sont atteintes 1 fois toutes les 50 charges (par construction du générateur de test). Il est donc normal que la convergence soit moins rapide dans le mode sinusoïdal de variation que dans le mode binaire.

Par exemple considérons le test 2.2. Cinquante charges sont produite en mode binaire de variation des teneurs.

Test 2.2						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C = 50, NMC = 7, Err = 0%						
Etapas 1 à 3 de l'algorithme						
m	IV(m)	MV(m)	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = .1$	IT(m,e) $\alpha = .2$	IT(m,e) $\alpha = .9$
1	85.1 - 100	B	85.106 - 100	84.279 - 100.827	83.245 - 101.862	18.085 - 167.022
2	85.2 - 100	B	85.487 - 100	84.681 - 100.806	83.673 - 101.814	20.178 - 165.309
3	85.3		85.3	85.3	85.3	85.3
4	85.4		85.356 - 85.400	85.353 - 85.402	85.350 - 85.406	85.156 - 85.600
5	85.5		85.457 - 85.500	85.454 - 85.502	85.451 - 85.505	85.262 - 85.695
6	85.6		85.6	85.6	85.6	85.6
7	85.7		85.7	85.7	85.7	85.7
8	85.8		85.8	85.8	85.8	85.8
9	85.9		85.853 - 85.900	85.850 - 85.903	85.847 - 85.906	85.641 - 86.112
10	86.0		85.929 - 86.000	85.925 - 86.004	85.920 - 86.009	85.608 - 86.320

La distance entre la solution trouvée et celle attendue n'est pas nulle. Elle est cependant relativement faible. Double le nombre de charges et regardons ce qu'il se passe.

Test 2.3						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C = 100, NMC = 7, Err = 0%						
Etapas 1 à 3 de l'algorithme						
m	IV(m)	MV(m)	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = .1$	IT(m,e) $\alpha = .2$	IT(m,e) $\alpha = .9$
1	85.1 - 100	B	85.1 - 100	84.272 - 100.828	83.237 - 101.863	18.050 - 167.050
2	85.2 - 100	B	85.2 - 100	84.378 - 100.822	83.350 - 101.850	18.600 - 166.600
3	85.3		85.3	85.3	85.3	85.3
4	85.4		85.4	85.4	85.4	85.4
5	85.5		85.5	85.5	85.5	85.5
6	85.6		85.6	85.6	85.6	85.6
7	85.7		85.7	85.7	85.7	85.7
8	85.8		85.8	85.8	85.8	85.8
9	85.9		85.9	85.9	85.9	85.9
10	86.0		86.0	86.0	86.0	86.0

Cette distance devient nulle. Quel que soit le nombre de charges supplémentaires produites cette distance reste nulle. Dans ce type de test, il existe donc un nombre minimum de charges à produire avec M types de mitraille au delà duquel il est certain que la distance

entre la solution trouvée et celle attendue vaille toujours 0. Une détermination empirique de ce nombre est envisageable.

Malheureusement il nous est impossible de pratiquer comme cela simplement parce que la puissance de calcul dont nous disposons est insuffisante. En effet, nous ne disposons que d'un pentium 100, et d'une routine FORTRAN non optimisée pour la résolution de simplexes. Les caractéristiques du compilateur dont nous disposons et/ou du système d'exploitation nous empêchent de réserver des tableaux de grosse taille. Nous ne sommes jamais arrivés à résoudre des PL de plus de 30 variables à 350 contraintes. Autrement dit, avec 10 types de mitraille, nous ne sommes jamais arrivés à simuler une production de plus de 150 charges.

Cependant, nous montrons que la production de 150 charges avec 10 types de mitraille et 7 types maximum de mitrilles par charge permet d'obtenir des distances faibles, voir nulle même en mode sinusoïdal.

Le cas le plus complexe que nous ayons simulé pour lequel cette distance valait toujours 0 est le suivant.

Test 2.4						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C = 100, NMC = 7, Err = 0%						
Etapes 1 à 3 de l'algorithme						
m	IV(m)	MV(m)	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = .1$	IT(m,e) $\alpha = .2$	IT(m,e) $\alpha = .9$
1	85.1 - 100	B	85.1 - 100	Ddct.	Ddct.	Ddct.
2	85.2 - 100	B	85.2 - 100	Ddct.	Ddct.	Ddct.
3	85.3 - 100	B	85.3 - 100	Ddct.	Ddct.	Ddct.
4	85.4 - 100	B	85.4 - 100	Ddct.	Ddct.	Ddct.
5	85.5 - 100	B	85.5 - 100	Ddct.	Ddct.	Ddct.
6	85.6	B	85.6	Ddct.	Ddct.	Ddct.
7	85.7	B	85.7	Ddct.	Ddct.	Ddct.
8	85.8	B	85.8	Ddct.	Ddct.	Ddct.
9	85.9	B	85.9	Ddct.	Ddct.	Ddct.
10	86.0	B	86.0	Ddct.	Ddct.	Ddct.

Ddct : Déductible

Pour des cas plus complexes (en mode sinusoïdal par exemple), il nous faudrait produire plus de 150 charges, pour que cette distance reste nulle. Malgré tout, avec 150 charges, cette distance reste « faible ». Par exemple, voici le test de type 2 le plus complexe que nous ayons réalisé

Test 2.5						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C= 150, NMC= 7, Err =0%						
m	IV(m)	MV(m)	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = .1$	IT(m,e) $\alpha = .2$	IT(m,e) $\alpha = .9$
1	70.1 - 100	S	70.490 - 97.297	Ddct.	Ddct.	Ddct.
2	70.2 - 100	S	71.567 - 100.353	Ddct.	Ddct.	Ddct.
3	70.3 - 100	S	72.712 - 100.037	Ddct.	Ddct.	Ddct.
4	70.4 - 100	S	72.027 - 100.873	Ddct.	Ddct.	Ddct.
5	70.5 - 100	S	71.224 - 99.090	Ddct.	Ddct.	Ddct.
6	80.6 - 90.6	S	80.790 - 89.592	Ddct.	Ddct.	Ddct.
7	80.7 - 90.7	S	81.699 - 90.677	Ddct.	Ddct.	Ddct.
8	80.8 - 90.8	S	82.591 - 88.904	Ddct.	Ddct.	Ddct.
9	80.9 - 90.9	S	81.185 - 90.904	Ddct.	Ddct.	Ddct.
10	81.0 - 91.0	S	82.093 - 89.249	Ddct.	Ddct.	Ddct.

La solution souhaitée et celle obtenue sont assez proches mais pas identiques. L'erreur maximale sur une des extrémités de l'intervalle trouvé est atteinte pour le premier type de mitraille. Elle vaut 2.7%. Cette solution nous permet malgré tout, de repérer correctement les tendances de variation des différents types de mitraille. Avec une production de 200 charges cette distance deviendrait peut-être nulle.

Rappelons que la production de 200 charges requiert en réalité 200 heures de travail, c.-à-d. à peu près 9 jours. Est-il industriellement possible de produire de l'acier pendant 9 jours non stop, avec 10 types de mitraille? Dans l'affirmative, si industriellement on respecte les hypothèses des tests de type 2, le modèle Sakawa-Yano a de bonnes chances de retrouver les caractéristiques de variation des teneurs. Ces hypothèses supposent une connaissance exacte des poids des types de mitraille utilisés et une probabilité égale du choix de chaque type de mitraille.

6.3.4 Les tests de type 3

Ces tests peuvent paraître inutiles dans la mesure où l'utilisation de régressions comme solutions au problème inverse est motivée par le fait que les différents types de mitraille ne sont pas homogènes. Or cette assertion est contraire aux hypothèses posées dans cette série de tests. Cependant, de manière similaire aux tests de type 2 nous nous focalisons ici sur l'effet sur la solution Sakawa-Yano des erreurs de type 2 uniquement. Nous essayerons de savoir comment le modèle Sakawa-Yano permet de tenir compte des erreurs de mesure des poids. Sakawa-Yano prétendent leur modèle plus général que ceux de Tanaka car il permettent de gérer des variables d'entrée (q(c,m)) floues.

Rappelons que le paramètre « Err » en entrée du générateur sert à simuler l'écart entre les poids mesurés et les poids réellement utilisés des différents types de mitraille.

Générons des données virtuelles de production ne contenant que des erreurs de type 2. Appliquons donc les 5 premières étapes de l'algorithme de génération avec une valeur de « Err » différente de 0 et des intervalles de variation des teneurs de longueurs non nulles. Les

données sont ainsi générées avec différents types de mitraille homogènes en fer et une différence entre les poids mesurés et les poids réellement utilisés.

L'introduction du flou dans la cinquième étape est censée permettre au modèle de tenir compte de cette approximation sur les quantités. Dans le test qui suit, les quantités mesurées peuvent varier de 5% par rapport aux quantités réellement utilisées. En introduisant un flou ($q_d(c,m)$) de 10% sur les quantités mesurées $q_c(c,m)$ des différents poids des types de mitraille, on espère que le modèle s'apercevra que les teneurs sont constantes. Autrement dit, on espère que le modèle n'attribuera pas l'écart à l'idéalité du réellement aux erreurs de type 2, en nous annonçant des erreurs de types 1.

Test 3.1											
Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano							
C=150,NMC = 7,Err = 5% Etapes 1 à 5 de l'algorithme				$\alpha = 0$		$\alpha = 0.1$		$\alpha = 0.9$		$\alpha = 0.9999$	
m	C(m)	R(m)	MV(m)	$t_c(m,e)$		$t_c(m,e)$		$t_c(m,e)$		$t_c(m,e)$	$t_d(m,e)$
1	85.1	0		80.259	0	80.995	0	85.149	22.622	85.171	31128.340
2	85.2	0		81.004	0	81.747	0	85.609	26.128	85.635	34680.700
3	85.3	0		80.835	0	81.576	0	85.424	26.144	85.450	34677.793
4	85.4	0		79.121	0	79.847	0	84.727	14.442	84.741	22905.849
5	85.5	0		83.404	0	84.169	0	87.865	29.707	87.895	38484.297
6	85.6	0		81.457	0	82.205	0	85.864	28.517	85.893	37095.031
7	85.7	0		79.715	0	80.446	0	85.566	12.519	85.579	21067.567
8	85.8	0		82.042	0	82.795	0	86.832	25.209	86.857	33883.113
9	85.9	0		79.505	0	80.235	0	84.861	17.293	84.878	25770.214
10	60.0	0		80.141	0	80.877	0	86.204	10.785	86.215	19396.960

Pour une valeur de α égale à 0 le modèle renseigne bien des teneurs constantes mais leur valeur n'est pas celle voulue. En réalité elle dépend uniquement du fait que l'on a mis un flou de 10% sur les variables d'entrée. Les valeurs obtenues et désirées ont donc aucun lien.

Pour une valeur de α égale à 1 il n'y a jamais de solution. En remplaçant $L^{-1}(\alpha)$ par la valeur 1 dans le PL résolu par Sakawa Yano, on constate que cela revient exactement à résoudre le même type de problème que dans les tests de type 2 avec les valeurs erronées des poids mesurés. Il est alors normal de ne pas trouver de solution.

Lorsque l'on augmente la valeur de α entre 0 et 1 le modèle garde ce genre de comportement. Il renseigne des teneurs constantes de plus en plus grandes toujours différentes des valeurs attendues. A partir de certaines valeurs de α , le modèle nous renseigne des teneurs non constantes ayant deux particularités. Les valeurs centrales des intervalles de variation trouvés semblent se rapprocher des valeurs désirées. Les rayons des intervalles de variation trouvés semblent devenir de plus en plus grands. L'introduction de flou sur les variables d'entrée est-il responsable de ce comportement ?

Dans le test 3.2, nous montrons que le résultat du test 3.1 n'est en rien meilleur à celui que l'on aurait obtenu en ne mettant pas de flou sur les valeurs mesurées des poids. Reprenons exactement les mêmes données de production que celles utilisées dans le test précédent et (ré)initialisons les valeurs de $q_d(c,m)$ à 0. Autrement dit, les données virtuelles

de production contiennent des erreurs de type 2 que l'on ne tente pas « d'absorber » en introduisant du flou. Il suffit en fait de se limiter aux quatre premières étapes de l'algorithme de simulation de production.

Test 3.2								
Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano				
C=150,NMC = 7,Err = 5% Etapes 1 à 4 de l'algorithme				$\alpha = 0$		$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
m	C(m)	R(m)	MV(m)	$t_e(m,e)$		$t(m,e)$	$t(m,e)$	$t(m,e)$
1	85.1	0		85.171	3.114	Ddct.	Ddct.	Ddct.
2	85.2	0		85.636	3.469	Ddct.	Ddct.	Ddct.
3	85.3	0		85.450	3.469	Ddct.	Ddct.	Ddct.
4	85.4	0		84.741	2.291	Ddct.	Ddct.	Ddct.
5	85.5	0		87.895	3.849	Ddct.	Ddct.	Ddct.
6	85.6	0		85.893	3.710	Ddct.	Ddct.	Ddct.
7	85.7	0		85.579	2.108	Ddct.	Ddct.	Ddct.
8	85.8	0		86.857	3.389	Ddct.	Ddct.	Ddct.
9	85.9	0		84.878	2.578	Ddct.	Ddct.	Ddct.
10	60.0	0		86.215	1.941	Ddct.	Ddct.	Ddct.

Les solutions trouvées dans les test 3.1 et 3.2 sont en fait liées. Cette relation est exprimable mathématiquement. Bien que nous ne sommes pas arrivés à le prouver, nous pensons avoir trouvé cette relation.

Soit $T_e^{s1}(\alpha)$ la solution Sakawa-Yano trouvée pour une valeur de α entre 0 et 1 à partir de données virtuelles de production générées avec les paramètres de simulation du type

$$\forall m \in \mathcal{M} \quad C(m) \in [0,100], R(m) = 0, \text{ étapes 1 à 5 de l'algorithme.}$$

Soit T_e^{s2} la solution Sakawa-Yano trouvée pour une valeur de α égale à 0 à partir des mêmes données virtuelles de production modifiées selon l'affectation

$$q_d(c,m) = 0 \quad \forall m \in \mathcal{M}$$

Alors on a que

$$\lim_{\alpha \rightarrow 1} T_e^{s1}(\alpha) = T_e^{s2}$$

$$\lim_{\alpha \rightarrow 1} T_d^{s1}(\alpha) / L^{-1}(\alpha) = T_d^{s2}$$

Nous pouvons illustrer cette relation à partir des résultats des deux tests 3.1 et 3.2. Considérons la teneur du type de mitraille 2. Etant donné que 0.9999 est très proche de 1.

$$\begin{aligned} 85.636 &\cong 85.635 \\ 34680.700 &\cong 3.469 / (1 - 0.9999) \end{aligned}$$

Nous en concluons donc que l'introduction de flou sur les variables d'entrée ($q(c,m)$) du modèle ne nous apporte rien. Que du contraire, elle nous est même défavorable en ce sens que les solutions du test 3.1 sont plus éloignées de la réalité que celles du test 3.2. Autrement dit, nous pouvons supprimer la cinquième étape de l'algorithme de base du générateur de test. C'est d'ailleurs ce que nous ferons lors de la quatrième série de tests.

Il est possible de comprendre ce phénomène. Les deux solutions des tests 3.1 et 3.2 introduisent toutes deux du flou sur les teneurs ($t_d(m,e) \neq 0$) alors que les données ont été générées sans variation des teneurs. Pour les tests de type 1 ce flou est une image de l'écart à l'idéalité (égalité $Q \otimes T_e = A_e$ dans les réels) due uniquement aux erreurs de type 2. Lors des tests de type 2 ce même flou est une image de l'écart à l'idéalité due uniquement aux erreurs de type 1. Or, il est impossible, à l'aide un seul paramètre (en l'occurrence T_e), même s'il est flou, de mettre en évidence lequel des deux types d'erreurs ne permet pas d'atteindre l'égalité $Q \otimes T_e = A_e$ dans les réels. Même si cette assertion n'est pour nous que du bon sens, nous imaginons que ce genre de résultat doit exister plus formellement dans la théorie de l'information. Remarquons qu'il en n'est pas du tout de même si A_e était un vecteur à composantes dans $F(\mathfrak{R})$ et pas dans \mathfrak{R} .

Nous décidons donc de ne plus mettre de flou sur les variables d'entrée. Ceci implique nécessairement que le modèle se trompe (test 3.2). Le modèle n'a aucun moyen de savoir si l'écart à l'idéalité est dû aux erreurs de type 2 ou aux erreurs de type 1. Nous avons vu que si il n'y avait pas d'imprécision sur les mesures des poids (pas d'erreurs de type 2), la solution trouvée serait exactement identique à celle attendue. Puisque l'on sait que les solutions trouvées sont d'autant plus altérées que les erreurs sur les poids mesurés sont grandes, il ne nous reste plus qu'à mesurer cette altération ?

Dans le test 3.2, l'erreur est inacceptable. Le modèle propose une teneur en fer variable dans l'intervalle $[85.171 - 3.114, 85.171 + 3.114] = [82.057, 88.285]$ alors qu'en réalité elle est constante et vaut 85.1% ! Une erreur de 5% sur les mesures des poids augmente beaucoup trop la distance entre la solution trouvée et celle attendue.

A titre d'ordre de grandeur, le test 3.3 est similaire au test 3.2. L'erreur de pesée est seulement de 1% ce qui est très (trop ?) raisonnable d'un point de vue industriel.

Test 3.3								
Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano				
C=150,NMC = 7,Err = 1% Etapas 1 à 4 de l'algorithme				$\alpha = 0$		$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
m	C(m)	R(m)	MV(m)	$t_c(m,e)$		$t(m,e)$	$t(m,e)$	$t(m,e)$
1	85.1	0		85.056	0.724	Ddct.	Ddct.	Ddct.
2	85.2	0		85.275	0.608	Ddct.	Ddct.	Ddct.
3	85.3	0		85.656	0.273	Ddct.	Ddct.	Ddct.
4	85.4	0		85.260	0.639	Ddct.	Ddct.	Ddct.
5	85.5	0		85.657	0.678	Ddct.	Ddct.	Ddct.
6	85.6	0		85.538	0.592	Ddct.	Ddct.	Ddct.
7	85.7	0		85.553	0.420	Ddct.	Ddct.	Ddct.
8	85.8	0		85.592	0.924	Ddct.	Ddct.	Ddct.
9	85.9	0		85.790	0.643	Ddct.	Ddct.	Ddct.
10	60.0	0		86.165	0.594	Ddct.	Ddct.	Ddct.

Les erreurs commises sont beaucoup plus petites que pour le test 3.2. Elles sont encore loin d'être négligeables.

Remarquons aussi que dans le test 3.2 la valeur 85.171 est assez proche de 85.1. Il s'agit « d'un hasard ». La même remarque est valable pour le test 3.3. C'est uniquement du au fait que nous avons généré des données de production pour lesquelles les poids mesurés sont, selon une distribution uniforme de probabilité, tantôt plus grand tantôt plus petits que les poids réellement utilisés. Il s'ensuit donc qu'en moyenne les valeurs mesurées et réellement utilisées sont proches. Si « par malheur » la mesure des poids est systématiquement plus grande de 5 % que les poids réellement utilisés, la valeurs 85.171 change et devient très différente de 85.1. C'est ce qui est réalisé dans le test 3.4.

Test 3.4								
Paramètres de simulation de production				Résultats de la méthode Sakawa-Yano				
C=150,NMC = 7,Err = 5% Etapas 1 à 4 de l'algorithme Systématiquement +5%				$\alpha = 0$		$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$
m	C(m)	R(m)	MV(m)	$t_c(m,e)$		$t(m,e)$	$t(m,e)$	$t(m,e)$
1	85.1	0		81.048	0	Ddct.	Ddct.	Ddct.
2	85.2	0		81.143	0	Ddct.	Ddct.	Ddct.
3	85.3	0		81.238	0	Ddct.	Ddct.	Ddct.
4	85.4	0		81.333	0	Ddct.	Ddct.	Ddct.
5	85.5	0		81.429	0	Ddct.	Ddct.	Ddct.
6	85.6	0		81.524	0	Ddct.	Ddct.	Ddct.
7	85.7	0		81.619	0	Ddct.	Ddct.	Ddct.
8	85.8	0		81.714	0	Ddct.	Ddct.	Ddct.
9	85.9	0		81.810	0	Ddct.	Ddct.	Ddct.
10	60.0	0		81.905	0	Ddct.	Ddct.	Ddct.

Les teneurs obtenues sont toutes systématiquement plus petites que celles attendues. Puisque l'on dit au modèle que toutes les quantités de mitraille utilisées sont plus grandes

qu'en réalité, pour une analyse en fer donnée (A_e) celui-ci en déduit que les teneurs en fer de tous les types de mitraille sont plus petites qu'en réalité.

6.3.5 Les tests de type 4

Dans cette série de test, les données virtuelles de productions contiennent les deux types d'erreurs. En mixant les observations des tests de type 1, 2, et 3, nous pouvons deviner le comportement du modèle.

Nous n'utilisons plus la dernière étape de l'algorithme de génération des tests. En effet, nous avons constaté que l'introduction de flou sur les variables d'entrée ($q(c,m)$) ne nous apporte rien.. Nous laissons au seul paramètre T_e (vecteur de nombre flous) expliquer l'écart à l'idéalité du aux deux types d'erreurs.

Nous nous limitons aussi à une valeur de Err = 1%. En effet, nous savons que pour des valeurs plus grandes la part du flou sur les variables indépendantes ($t(m,e)$) due aux écart de type 1 est fortement altérée par celle due aux erreurs de type 2 (quantifiées par « Err »). Pour une valeur de 1% cette altération devient plus raisonnable. Il est alors encore possible d'obtenir une approximation de la variation des teneurs en fer des types de mitraille.

Par exemple, reprenons le test 2.4 mais introduisons une erreur de 1% sur les poids mesurés.

Test 4.1						
Paramètres de simulation de production			Résultats de la méthode Sakawa-Yano			
C = 150, NMC = 7, Err = 1% Etapas 1 à 4 de l'algorithme						
m	IV(m)	MV(m))	IT(m,e) $\alpha = 0$	IT(m,e) $\alpha = .1$	IT(m,e) $\alpha = .2$	IT(m,e) $\alpha = .9$
1	85.1 - 100	B	84.387 - 100.418	Ddct.	Ddct.	Ddct.
2	85.2 - 100	B	85.472 - 99.966	Ddct.	Ddct.	Ddct.
3	85.3 - 100	B	84.602 - 101.045	Ddct.	Ddct.	Ddct.
4	85.4 - 100	B	85.309 - 100.188	Ddct.	Ddct.	Ddct.
5	85.5 - 100	B	84.812 - 100.667	Ddct.	Ddct.	Ddct.
6	85.6	B	85.365 - 85.716	Ddct.	Ddct.	Ddct.
7	85.7	B	85.303 - 86.183	Ddct.	Ddct.	Ddct.
8	85.8	B	84.957 - 86.045	Ddct.	Ddct.	Ddct.
9	85.9	B	85.261 - 87.044	Ddct.	Ddct.	Ddct.
10	86.0	B	85.539 - 86.612	Ddct.	Ddct.	Ddct.

La distance entre la solution obtenue et celle désirée n'est plus nulle. Elle l'était dans le test 2.4. Cette solution nous permet de repérer une très forte variation de la teneur en fer pour les cinq premiers types de mitraille et une plus faible pour les 5 derniers. Nous ne pouvons malheureusement plus retrouver les caractéristiques exactes de variation.

6.4 La compréhension des résultats

Nous essayons dans cette partie de comprendre mathématiquement le comportement du modèle. Nous reprenons pour cela chacune des étapes de test.

6.4.1 Les tests de type 1

Dans ce type de test, des données virtuelles de production ne contiennent aucun des deux types d'erreur. Mathématiquement cela s'exprime par les propriétés suivantes.

1. On sait que $q_c(c,m)$ vaut exactement le poids réellement utilisé du type de mitraille m dans la charge c . Il n'est donc pas nécessaire d'introduire du flou sur ce paramètre.

$$q_d(c,m) = 0 \quad \forall m \in \mathcal{M} \quad \forall c \in \mathcal{C}.$$

2. Les mitrailles sont homogènes en fer (élément e). L'égalité $Q \otimes T_e = A_e$ peut donc être respectée dans les réels. Soit $t^s(m,e) \in \mathcal{R}$ la teneur constante en fer du type de mitraille M . On a donc que

$$\sum_{m \in \mathcal{M}} q_c(c,m) t^s(m,e) = a(e,c) \quad \forall c \in \mathcal{C}$$

Suite à (1), le PL résolu par le modèle Sakawa-Yano est fortement simplifié. Il devient le suivant.

$$\begin{aligned} \text{MIN} \quad \text{FO2} &= \sum_{m \in \mathcal{M}} \left[\left(2 \sum_{c \in \mathcal{C}} q_c(c,m) \right) t_d(m,e) \right] \\ \text{SC} \quad \sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) - L^{-1}(\alpha) t_d(m,e)) &\leq a(e,c) \quad \forall c \in \mathcal{C} \\ \sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) + L^{-1}(\alpha) t_d(m,e)) &\geq a(e,c) \quad \forall c \in \mathcal{C} \\ t_c(m,e) - L^{-1}(\alpha) t_d(m,e) &\geq 0 \quad \forall m \in \mathcal{M} \\ t_c(m,e) &\geq 0 \quad \forall m \in \mathcal{M} \end{aligned}$$

Suite à (2), il est évident que quelle que soit la valeur de α la solution est obtenue pour une valeur nulle de FO2 et donc que

$$\begin{aligned} t_d(m,e) &= 0 \quad \forall m \in \mathcal{M} \\ t_c(m,e) &= t^s(m,e) \quad \forall m \in \mathcal{M} \end{aligned}$$

6.4.2 Les tests de type 2

Dans ce type de test, les données virtuelles de production contiennent uniquement des erreurs de types 1. Nous avons remarqué qu'une augmentation de α ne modifie pas les centres des solutions, et augmente le flou de façon linéaire. Nous prouvons ici cette relation de linéarité.

- On sait que $q_c(c,m)$ vaut exactement le poids réellement utilisé du type de mitraille m dans la charge c . Il n'est pas nécessaire d'introduire du flou sur ce paramètre.

$$q_d(c,m) = 0 \quad \forall m \in \mathcal{M} \quad \forall c \in \mathcal{C}.$$

Le PL résolu par le modèle Sakawa-Yano est donc fortement simplifié. Il a exactement la même forme que pour les tests de type 1.

$$\text{MIN} \quad \text{FO2} = \sum_{m \in \mathcal{M}} \left[\left(2 \sum_{c \in \mathcal{C}} q_c(c,m) \right) t_d(m,e) \right]$$

$$\text{SC} \quad \sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) - L^{-1}(\alpha) t_d(m,e)) \leq a(e,c) \quad \forall c \in \mathcal{C}$$

$$\sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) + L^{-1}(\alpha) t_d(m,e)) \geq a(e,c) \quad \forall c \in \mathcal{C}$$

$$t_c(m,e) - L^{-1}(\alpha) t_d(m,e) \geq 0 \quad \forall m \in \mathcal{M}$$

$$t_c(m,e) \geq 0 \quad \forall m \in \mathcal{M}$$

Contrairement aux tests de type 1, l'égalité $Q \otimes T_e = A_e$ ne peut plus être respectée car les types de mitraille ne sont plus homogènes.

Nous montrons ici que les solutions obtenues pour deux valeurs différentes de α sont déductibles l'une de l'autre. En effet, suite à une petite manipulation mathématique on peut exprimer le PL d'une autre manière

$$\text{MIN} \quad \text{FO2} = \sum_{m \in \mathcal{M}} \left[\left(2 \sum_{c \in \mathcal{C}} q_c(c,m) \right) t_d(m,e) \right]$$

$$\text{SC} \quad L^{-1}(\alpha) \left(\sum_{m \in \mathcal{M}} q_c(c,m) t_d(m,e) \right) \geq \left| \sum_{m \in \mathcal{M}} q_c(c,m) t_c(m,e) - a(e,c) \right| \quad \forall c \in \mathcal{C}$$

$$t_c(m,e) - L^{-1}(\alpha) t_d(m,e) \geq 0 \quad \forall m \in \mathcal{M}$$

$$t_c(m,e) \geq 0 \quad \forall m \in \mathcal{M}$$

Soit $(t_c^1(m,e), t_d^1(m,e))$ où $m \in \mathcal{M}$, la solution trouvée pour α_1

Soit $(t_c^2(m,e), t_d^2(m,e))$ où $m \in \mathcal{M}$, la solution trouvée pour $\alpha_2 \neq \alpha_1$

En fonction de l'allure du PL à résoudre, on en déduit que

$$t_c^1(m,e) = t_c^2(m,e) \quad \forall m \in \mathcal{M}$$

$$t_d^2(m,e) = (L^{-1}(\alpha_1)/L^{-1}(\alpha_2)) t_d^1(m,e)$$

Nous ne sommes pas arrivés à prouver en toute généralité que dans les conditions de réalisation de ces tests de type 2, la solution trouvée converge (par augmentation du nombre de charges) vers la solution désirée. En interprétant géométriquement et physiquement le modèle Sakawa-Yano à partir de la formulation du PL donnée ci dessus (en terme de valeur

absolue), nous pouvons comprendre intuitivement pourquoi il en est ainsi. Nous n'aborderons pas ici ces considérations géométriques car elles s'adaptent plus particulièrement à un exposé oral.

6.4.3 Les tests de type 3 et 4

De façon similaire à précédemment, nous disposons d'arguments géométriques pour expliquer le comportement du modèle. Etant donné que nous n'avons encore jamais introduit d'interprétation géométrique du modèle Sakawa-Yano, nous les réservons pour la présentation orale.

6.5 La solution finale proposée

En fonction de tous les tests que nous avons réalisés, nous proposons de modifier la solution générale Sakawa-Yano. Nous pouvons y introduire deux simplifications majeures

- $q_d(c,m) = 0 \quad \forall m \in \mathcal{M} \forall c \in \mathcal{C}$
- $\alpha = 0$

Le problème à résoudre est donc le suivant.

$$\begin{array}{ll}
 \forall e \in \mathcal{E} \\
 \text{MIN} & \text{FO2} = \sum_{m \in \mathcal{M}} \left[\left(2 \sum_{c \in \mathcal{C}} q_d(c,m) \right) t_c(m,e) + \left(2 \sum_{c \in \mathcal{C}} q_c(c,m) \right) t_d(m,e) \right] \\
 \text{SC} & \sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) - t_d(m,e)) \leq a(e,c) \quad \forall c \in \mathcal{C} \\
 & \sum_{m \in \mathcal{M}} q_c(c,m) (t_c(m,e) + t_d(m,e)) \geq a(e,c) \quad \forall c \in \mathcal{C} \\
 & t_c(m,e) - t_d(m,e) \geq 0 \quad \forall m \in \mathcal{M} \\
 & t_c(m,e) \geq 0 \quad \forall m \in \mathcal{M}
 \end{array}$$

Nous évitons donc deux problèmes importants. D'une part, nous ne devons plus essayer de balayer les optimum de Pareto comme le propose les auteurs. D'autre part, nous ne devons pas introduire de flou sur nos variables d'entrée. Le fait que nos variables de sortie ($a(e,c)$) ne soient pas floues, c.-à-d. que les analyses réalisées en sortie du circuit de production soient connues exactement et donc ne sont pas « floues », en est la raison principale.

Lorsque $q_c(c,m)$ représente exactement la quantité du type de mitraille m utilisée dans la charge c et qu'on a produit un assez grand nombre de charges (C assez grand), la solution trouvée $[t_c(m,e) - t_d(m,e), t_c(m,e) + t_d(m,e)]$ est l'intervalle dans lequel la teneur en l'élément e du type de mitraille m a varié pour la production des C charges.

Lorsque $q_c(c,m)$ est toujours « assez proche » de la quantité du type de mitraille m utilisée dans la charge c et qu'on a produit un assez grand nombre de charges (C assez grand), la solution trouvée $[t_c(m,e) - t_d(m,e), t_c(m,e) + t_d(m,e)]$ est une approximation de l'intervalle dans lequel la teneur en l'élément e du type de mitraille m a varié pour la production des C charges.

Il serait intéressant de réaliser des tests à l'échelle industrielle pour déterminer ce que l'on entend par « assez proche » et « assez grand ».

Chapitres 7 et 8

Conclusion générale et bibliographie

7. Conclusion générale

L'objectif de ce mémoire était d'étudier la théorie des ensembles flous en vue de l'élaboration d'une solution meilleure à un des trois problèmes de production proposés par profilARBED. Une préférence avait été marquée pour le problème inverse car industriellement il est de loin le plus intéressant à résoudre. Il y avait là un défi extrêmement tentant à relever.

Après de multiples recherches qu'il nous fallait canaliser de manière à aller droit au but, nous nous sommes effectivement dirigés vers ce problème inverse. C'est ainsi que l'étude de la théorie des ensembles flous s'est principalement déroulée autour de la compréhension du modèle de régression linéaire floue proposé par Sakawa et Yano.

La notion même de degré d'appartenance sur lequel se base la théorie des ensembles flous est très subjective. Elle peut être interprétée de multiples manières. Intrinsèquement, elle laisse beaucoup de liberté quant à la manière dont on peut développer et utiliser cette théorie. Il est donc légitime de se demander si ce n'est pas une porte ouverte à des scientifiques peu scrupuleux ? Toutefois, certains d'entre eux se sont penchés sur le problème avec sérieux. Lorsque Zadeh propose sa vue particulière de la théorie des ensembles flous en reliant la notion de degré d'appartenance à la notion de possibilité, nous disposons là d'un nouvel outil intéressant. Ce que nous avons appelé la théorie possibiliste des ensembles flous est à la base du développement d'outils efficaces utilisés dans des domaines manipulant des connaissances imprécises. Cet engouement en est presque devenu une mode et il ne tient qu'à notre esprit critique de faire la part des choses.

Quoi qu'il en soit, il est indéniable que la méthode de régression linéaire floue proposée par Sakawa et Yano nous a permis de développer une solution meilleure au problème inverse que celle qui existait auparavant. Lorsque les mesures industrielles dont on dispose sont connues précisément, cette méthode permet de retrouver exactement les caractéristiques de variation des teneurs des types de mitrailles. Lorsque ces mêmes données sont connues avec une bonne précision, c.-à.-d. 1% d'erreur maximum, cette méthode permet d'avoir une approximation de ces variations. Lorsque l'imprécision est trop grande il devient « impossible » de retrouver ces caractéristiques de variation. Mais cette constatation n'est elle pas rassurante ?

Nous espérons donc que cette petite étude en collaboration avec l'ARBED, permettra de développer un bon outil industriel pour la production d'acier. Dans l'état actuel des choses, il y a très probablement moyen de développer un bon outil d'aide à la décision sur base de cette méthode. De plus, en y ajoutant certaines heuristiques de production et en restant dans les limites du raisonnable quant à la connaissance des données industrielles, rien ne dit qu'il n'est pas possible de l'améliorer.

8. Bibliographie

8.1 Introduction à la théorie des ensembles flous

[BM] Bernadette Bouchon-Meunier, « La logique floue », Que sais-je ?, Presses Universitaires de France, ISBN 2 13 045007 5

[Nguyen78] Hung T. Nguyen, « A note on the extension Principle for Fuzzy Sets », Journal of mathematical analysis and applications, 64 (1978), 369-380

[HJZ93] H.-J. Zimmermann, « Fuzzy Sets and its applications », Prentice Hall, 1993

8.2 Le problème direct

[DVV] Delgado, J.L. Verdegay, M.A. Vila, « Relating different approaches to solve linear programming problems with fuzzy constraints » (avril 1989), Fuzzy Sets and systems, ?, 33-42

[HVZ] F. Herrera, J.L Verdegay, H.-J. Zimmermann, « Boolean programming problems with fuzzy constraints » (Septembre 1992), Fuzzy Sets and systems, 55 (1993), 285-293

8.3 Le problème inverse

[Bard90] A. Bárdossy, « Note on fuzzy regression » (February 1989), Fuzzy Sets and systems, 35 (1990), 65-75

[HT89] I. Hayashi, H Tanaka, « The fuzzy gmdh algorithm by possibilist models and its application » (January 1989), Fuzzy Sets and systems, 36 (1990), 245-258

[HK85] B. Heshmaty, A. Kandel, « Fuzzy linear regression and its applications to forecasting in uncertain environment » (March 1984), Fuzzy Sets and systems, 15 (1985), 159-191

[KC] S. Kundu, J. Chen, « FLIC : Fuzzy Linear Invariant Clustering For Applications In Fuzzy Control », Computer Science department, Louisiana State university, Baton Rouge, LA 70803.

[MK93] H. Moskowitz, Kwangjae Kim, « On assessing the H value in fuzzy linear regression » (February 1993), Fuzzy Sets and systems, 58 (1993), 303-327

[OKL88] S.-B. Oh, W. Kim, J.K. Lee, « An approach to causal modeling in fuzzy environment and its application » (August 1988), Fuzzy Sets and systems, 35 (1990), 43-55

[Pet93] G. Peters, « Fuzzy linear regression with fuzzy intervals » (December 1993), Fuzzy Sets and systems, 63 (1993), 45-55

[SS94] M. Sakawa, K. Sawada, « An interactive fuzzy satisficing method for large-scale multiobjective linear programming problems with block angular structure », Fuzzy Sets and systems, 67 (1994), 5-17

[SY92] M. Sakawa, H Yano, « Multiobjective fuzzy linear regression analysis for fuzzy input-output data » (September 1990), Fuzzy Sets and systems, 47 (1992), 19-27

[SY94] M. Sakawa, H Yano, « A fuzzy dual decomposition method for large-scale multiobjective nonlinear programming problems » (December 1994), Fuzzy Sets and systems, 67 (1994), 19-27

[Tan87] H. Tanaka, « Fuzzy data analysis by possibilistic linear models » (june 1986), Fuzzy Sets and systems, 24 (1987), 363-375

[THW89] H. Tanaka, I. Hayashi, J. Watada, « Possibilistic linear regression analysis for fuzzy data », European Journal of Operational Research, 40 (1989), 389-396

[TW86] H. Tanaka, J. Watada, « Possibilistic linear systems and their application to the linear regression model » (December 1896), Fuzzy Sets and systems, 27 (1988), 275-289

[WKS93] L. Richard Woodyatt, Kenneth L. Stott, Floyd E. Wolf, Francis J. Vasko, « An application combining set covering and fuzzy sets to optimally assign metallurgical grades to customer orders » (1992), Fuzzy Sets and systems, 53 (1993), 15-25

8.4 Le problème de planification de production

[BF94] K.P. Bernatzki, D. Fengler, H.P. Kaiser, Wolf Lanzer, « Disposition und Zeitwirtschaft in Stahlwerk Huckingen unter Einsatz von Fuzzy-Technologie », Stahl und Eisen 114 (1994), Nr. 5, 89-95

[LL94] W. Limbeck, H.P. Lipp, « Ein fuzzy-konzept zur operativen Führung des Stahlwerksprozesses », Stahl und Eisen 114 (1994), 5, 81-87

[Lipp94] H.-P. Lipp, « Prinzipien und Potentiale de Fuzzy-Technologien in der Automatisierungstechnik », Preussag, Informations und Kooperationskreis Physikalische Techniken, 21 Juni 1994

[VAC94] R. Valette, D. Andreu, J. Cardoso, J.-C. Pascal, « Fuzzy Petri Nets and their Application in CIME », T.IEE Japan. 114-C (1994), 9

Annexe

Multiobjective fuzzy linear regression analysis for fuzzy input–output data

Masatoshi Sakawa

Department of Industrial and Systems Engineering, Faculty of Engineering, Hiroshima University, Higashi-Hiroshima 724, Japan

Hitoshi Yano

Nagoya Municipal Women's College, Nagoya 464, Japan

Received August 1989

Revised September 1990

Abstract: Fuzzy linear regression models, where both input data and output data are fuzzy numbers, are introduced by using three indices for equalities between fuzzy numbers. By considering the conflict between the fuzzy threshold for the three indices and the fuzziness of the fuzzy linear regression model, three types of multiobjective programming problems for obtaining fuzzy linear regression models are formulated corresponding to the three indices. Then a linear programming based interactive decision making method to derive the satisficing solution of the decision maker for the formulated multiobjective programming problems is developed. A numerical example demonstrates the appropriateness and efficiency of the proposed method.

Keywords: Fuzzy linear regression; possibility; necessity; fuzzy input–output data; multicriteria analysis.

1. Introduction

In fuzzy linear regression analysis, recently proposed by Tanaka et al. [9–11], deviations between the observed values and the estimated values are assumed to depend on the fuzziness of the system structure, in contrast to the usual linear regression analysis where deviations are supposed to be caused by observation errors or variables not in the model. Linear programming based methods for obtaining fuzzy linear regression models for some fixed fuzzy threshold via the set inclusion relations among α -level sets for fuzzy numbers, or possibility and necessity measures have been proposed [7, 9–11].

More recently, by considering the conflict

between the fuzzy threshold for the three indices for equalities between fuzzy numbers [2, 4] and the fuzziness of the fuzzy linear regression model, Sakawa and Yano [8] formulated three types of multiobjective programming problems for obtaining fuzzy linear regression models, and developed a linear programming based interactive decision making method to derive the satisficing solution of the decision maker.

In these fuzzy linear regression models [7, 8, 9–11], however, it should be stressed that while output data is assumed to be a fuzzy number, input data is not a fuzzy number.

In this paper, as a more generalized version of Sakawa and Yano [7, 8], three types of multiobjective programming problems for obtaining fuzzy linear regression models, where both input data and output data are fuzzy numbers, are formulated by using three indices for equalities between two fuzzy numbers [2, 4], and an interactive decision making method to derive the satisficing solution of the decision maker is developed on the basis of the linear programming method.

2. Fuzzy linear regression models

In this paper, we consider fuzzy linear regression models where both input data and output data are represented as fuzzy numbers. To cope with such fuzzy input–output data, we assume the fuzzy linear regression model is represented as follows:

$$\bar{y}_i = \bar{a} \otimes \bar{x}_i, \quad i = 1, \dots, k, \quad (1)$$

where

$$\bar{a} = (\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n), \quad \bar{x}_i = (1, \bar{x}_{i1}, \dots, \bar{x}_{in})^T, \quad (2)$$

and \bar{y}_i , \bar{a}_j , and \bar{x}_{ij} are fuzzy numbers. Observe that ' \otimes ' is the product operator based on the extension principle [12].

In the following, for simplicity, we assume

that the fuzzy input data \tilde{x}_{ij} , the fuzzy output data \tilde{y}_i , and the fuzzy parameters \tilde{a}_j are given as L-L fuzzy numbers [3] defined by

$$\mu_{\tilde{x}_{ij}}(x_{ij}) = L(|\tilde{x}_{ij} - x_{ij}|/d_{ij}), \quad d_{ij} > 0, \quad (3)$$

$$\mu_{\tilde{y}_i}(y_i) = L(|\tilde{y}_i - y_i|/e_i), \quad e_i > 0, \quad (4)$$

$$\mu_{\tilde{a}_j}(a_j) = L(|\tilde{a}_j - a_j|/c_j), \quad c_j > 0, \quad (5)$$

where $L(\cdot)$ is a reference function, \tilde{x}_{ij} , \tilde{y}_i and \tilde{a}_j are mean values, and d_{ij} , e_i and c_j are spread parameters [3]. When spread parameters are zero, i.e., $d_{ij} = 0$, $e_i = 0$ and $c_j = 0$, \tilde{x}_{ij} , \tilde{y}_i and \tilde{a}_j are nonfuzzy numbers. For such nonfuzzy numbers \tilde{x}_{ij} , \tilde{y}_i and \tilde{a}_j with zero spread parameters, we define the corresponding membership functions $\mu_{\tilde{x}_{ij}}(x_{ij})$, $\mu_{\tilde{y}_i}(y_i)$ and $\mu_{\tilde{a}_j}(a_j)$ as follows:

$$\mu_{\tilde{x}_{ij}}(x_{ij}) = \begin{cases} 1, & x_{ij} = \tilde{x}_{ij}, \\ 0, & x_{ij} \neq \tilde{x}_{ij}, \end{cases}$$

$$\mu_{\tilde{y}_i}(y_i) = \begin{cases} 1, & y_i = \tilde{y}_i, \\ 0, & y_i \neq \tilde{y}_i, \end{cases}$$

$$\mu_{\tilde{a}_j}(a_j) = \begin{cases} 1, & a_j = \tilde{a}_j, \\ 0, & a_j \neq \tilde{a}_j. \end{cases}$$

Such L-L fuzzy numbers \tilde{x}_{ij} , \tilde{y}_i and \tilde{a}_j are often symbolically represented by [3]

$$\tilde{x}_{ij} = (\tilde{x}_{ij}, d_{ij})_L, \quad (6)$$

$$\tilde{y}_i = (\tilde{y}_i, e_i)_L, \quad \tilde{a}_j = (\tilde{a}_j, c_j)_L.$$

For notational convenience, denote the vector of the fuzzy parameters $\tilde{a} = (\tilde{a}_0, \dots, \tilde{a}_n)$ by

$$\tilde{a} = (\tilde{a}, c)_L \quad (7)$$

where

$$\tilde{a} = (\tilde{a}_0, \tilde{a}_1, \dots, \tilde{a}_n), \quad c = (c_0, c_1, \dots, c_n). \quad (8)$$

Furthermore, it is also assumed that the α -level set of the fuzzy number \tilde{x}_{ij} is nonnegative for any $\alpha \in (0, 1]$.

It should be emphasized here that although the explicit form of the membership function of $\tilde{a} \otimes \tilde{x}_i$ can not directly be obtained, the α -level set of $\tilde{a} \otimes \tilde{x}_i$ can be obtained from the result of Nguyen [6] as follows:

Theorem 1.

$$(\tilde{a} \otimes \tilde{x}_i)_\alpha = [Y_{i\alpha}^L, Y_{i\alpha}^R] \quad (9)$$

where

$$Y_{i\alpha}^L = \sum_{j \in J_1 \cup J_2} \{ \min \{ (\tilde{a}_j - L^{-1}(\alpha)c_j)(\tilde{x}_{ij} - L^{-1}(\alpha)d_{ij}), (\tilde{a}_j - L^{-1}(\alpha)c_j)(\tilde{x}_{ij} + L^{-1}(\alpha)d_{ij}) \} \} \quad (10)$$

$$Y_{i\alpha}^R = \sum_{j \in J_3} \{ \max \{ (\tilde{a}_j + L^{-1}(\alpha)c_j)(\tilde{x}_{ij} + L^{-1}(\alpha)d_{ij}), (\tilde{a}_j + L^{-1}(\alpha)c_j)(\tilde{x}_{ij} - L^{-1}(\alpha)d_{ij}) \} \} \quad (11)$$

It should be noted here that $Y_{i\alpha}^L$ and $Y_{i\alpha}^R$, $i = 1, \dots, k$, involve the minimization and maximization operators.

In order to deal with these operators in $Y_{i\alpha}^L$ and $Y_{i\alpha}^R$, $i = 1, \dots, k$, assume that the following relations hold for any fixed degree α :

$$\tilde{a}_j - L^{-1}(\alpha)c_j \geq 0, \quad j \in J_1, \quad (12)$$

$$\tilde{a}_j - L^{-1}(\alpha)c_j \leq 0, \quad \tilde{a}_j + L^{-1}(\alpha)c_j \geq 0, \quad j \in J_2, \quad (13)$$

$$\tilde{a}_j + L^{-1}(\alpha)c_j \leq 0, \quad j \in J_3, \quad (14)$$

where

$$J = \{0, \dots, n\} = J_1 \cup J_2 \cup J_3,$$

$$J_i \cap J_j = \emptyset, \quad i(\neq j) = 1, 2, 3.$$

Such assumptions may not be so restrictive since, in real world situations, the sign of parameters of the linear regression model is often known in advance. On the basis of assumptions, let us introduce the following which depends on the index sets J_1 , J_2 and J_3 :

$$L(J_1, J_2, J_3) = \{(\tilde{a}, c) \in R^{2(n+1)} \mid \begin{aligned} &\tilde{a}_j - L^{-1}(\alpha)c_j \geq 0, \quad c_j \geq 0, \quad j \in J_1, \\ &\tilde{a}_j - L^{-1}(\alpha)c_j \leq 0, \quad \tilde{a}_j + L^{-1}(\alpha)c_j \geq 0, \quad j \in J_2, \\ &\tilde{a}_j + L^{-1}(\alpha)c_j \leq 0, \quad c_j \geq 0, \quad j \in J_3 \end{aligned} \}$$

Then $Y_{i\alpha}^L$ and $Y_{i\alpha}^R$, $i = 1, \dots, k$, can be expressed as

$$Y_{i\alpha}^L = \sum_{j \in J_1} (\tilde{a}_j - L^{-1}(\alpha)c_j)(\tilde{x}_{ij} - L^{-1}(\alpha)d_{ij}) + \sum_{j \in J_2 \cup J_3} (\tilde{a}_j - L^{-1}(\alpha)c_j)(\tilde{x}_{ij} + L^{-1}(\alpha)d_{ij})$$

$$Y_{i\alpha}^R = \sum_{j \in J_1 \cup J_2} (\tilde{a}_j + L^{-1}(\alpha)c_j)(\tilde{x}_{ij} + L^{-1}(\alpha)d_{ij}) + \sum_{j \in J_3} (\tilde{a}_j + L^{-1}(\alpha)c_j)(\tilde{x}_{ij} - L^{-1}(\alpha)d_{ij}). \quad (16)$$

To cope with the fuzzy equality in the fuzzy linear regression model (1), we now introduce the following three indices for the equalities between two fuzzy numbers \tilde{m} and \tilde{n} , as proposed by Dubois and Prade [2, 4]:

$$\text{Pos}(\tilde{m} = \tilde{n}) = \sup_{u \in R^1} \min \{ \mu_{\tilde{m}}(u), \mu_{\tilde{n}}(u) \}, \quad (17)$$

$$\text{Nes}(\tilde{m} \subset \tilde{n}) = \inf_{u \in R^1} \max \{ 1 - \mu_{\tilde{m}}(u), \mu_{\tilde{n}}(u) \}, \quad (18)$$

$$\text{Nes}(\tilde{m} \supset \tilde{n}) = \inf_{u \in R^1} \max \{ \mu_{\tilde{m}}(u), 1 - \mu_{\tilde{n}}(u) \}. \quad (19)$$

where $\mu_{\tilde{m}}(u)$ and $\mu_{\tilde{n}}(u)$ are the membership functions of \tilde{m} and \tilde{n} . and Pos and Nes are short for Possibility and Necessity.

By adopting these three indices for the equality relation in the fuzzy linear regression model, three types of problems for obtaining the fuzzy linear regression models where both input data and output data are fuzzy numbers, can be constructed [7, 11]:

$$P_1(\alpha): \min Z = \sum_{i=1}^k (Y_{i\alpha}^R - Y_{i\alpha}^L) \quad (20)$$

subject to

$$\text{Pos}(\tilde{y}_i = \tilde{a} \otimes \tilde{x}_i) \geq \alpha, \quad i = 1, \dots, k,$$

$$(\tilde{a}, c) \in L(J_1, J_2, J_3).$$

$$P_2(\alpha): \min Z = \sum_{i=1}^k (Y_{i\alpha}^R - Y_{i\alpha}^L) \quad (21)$$

subject to

$$\text{Nes}(\tilde{y}_i \subset \tilde{a} \otimes \tilde{x}_i) \geq \alpha, \quad i = 1, \dots, k,$$

$$(\tilde{a}, c) \in L(J_1, J_2, J_3),$$

$$P_3(\alpha): \max Z = \sum_{i=1}^k (Y_{i\alpha}^R - Y_{i\alpha}^L) \quad (22)$$

subject to

$$\text{Nes}(\tilde{y}_i \supset \tilde{a} \otimes \tilde{x}_i) \geq \alpha, \quad i = 1, \dots, k,$$

$$(\tilde{a}, c) \in L(J_1, J_2, J_3),$$

where the value α ($0 < \alpha \leq 1$) represents the degree of conformity between the fuzzy output data \tilde{y}_i and the fuzzy linear regression model

$\tilde{a} \otimes \tilde{x}_i$, and the objective function Z means the degree of the fuzziness of the fuzzy linear regression model (1).

From the properties of the three indices [2, 7] and the definition of L-L fuzzy numbers \tilde{y}_i , \tilde{x}_{ij} , \tilde{a}_j , the inequalities of the problems $P_1(\alpha)$, $P_2(\alpha)$ and $P_3(\alpha)$ can be transformed to the usual inequalities.

Theorem 2. (1) $\text{Pos}(\tilde{y}_i = \tilde{a} \otimes \tilde{x}_i) \geq \alpha$, $i = 1, \dots, k$, if and only if

$$\tilde{y}_i + L^{-1}(\alpha)e_i \geq Y_{i\alpha}^L, \quad \tilde{y}_i - L^{-1}(\alpha)e_i \leq Y_{i\alpha}^R. \quad (23)$$

(2) $\text{Nes}(\tilde{y}_i \subset \tilde{a} \otimes \tilde{x}_i) \geq \alpha$, $i = 1, \dots, k$, if and only if

$$\tilde{y}_i + L^{-1}(1 - \alpha)e_i \leq Y_{i\alpha}^R, \quad \tilde{y}_i - L^{-1}(1 - \alpha)e_i \geq Y_{i\alpha}^L. \quad (24)$$

(2) $\text{Nes}(\tilde{y}_i \supset \tilde{a} \otimes \tilde{x}_i) \geq \alpha$, $i = 1, \dots, k$, if and only if

$$\tilde{y}_i + L^{-1}(\alpha)e_i \geq Y_{i,1-\alpha}^R, \quad \tilde{y}_i - L^{-1}(\alpha)e_i \leq Y_{i,1-\alpha}^L. \quad (25)$$

It should be noted here that, for some fixed degree α , the inequalities (23), (24) and (25) are linear with respect to (\tilde{a}, c) , since $Y_{i\alpha}^L$ and $Y_{i\alpha}^R$ are linear functions with respect to (\tilde{a}, c) as shown in (15) and (16). From this fact, the problems $P_1(\alpha)$, $P_2(\alpha)$ and $P_3(\alpha)$ can be reduced to linear programming problems with respect to (\tilde{a}, c) . Therefore, it is easy to solve these problems by using the linear programming method.

However, before solving the linear programming problems $P_1(\alpha)$, $P_2(\alpha)$ and $P_3(\alpha)$, the value of the degree α must be selected within the interval $(0, 1]$. In general, a higher degree α of the fuzzy threshold is more preferable because α can be interpreted as the degree of the equivalence between the fuzzy output data \tilde{y}_i and the estimated value $\tilde{a} \otimes \tilde{x}_i$ of the fuzzy linear regression model (1).

From this point of view, by considering the conflict between maximizing the degree α of the fuzzy threshold, and minimizing (or maximizing) the fuzziness Z of the fuzzy linear regression model in the problems $P_1(\alpha)$ and $P_2(\alpha)$ (or $P_3(\alpha)$), we can formulate the following multiobjective programming problems for ob-

taining the fuzzy linear regression models where both input data and output data are fuzzy numbers:

$$\text{MP}_1: \min Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L), \quad (26)$$

max α
subject to

$$\text{Pos}(\bar{y}_i = \bar{a} \otimes \bar{x}_i) \geq \alpha, \quad i = 1, \dots, k, \\ (\bar{a}, c) \in L(J_1, J_2, J_3), \quad 0 < \alpha \leq 1.$$

$$\text{MP}_2: \min Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L), \quad (27)$$

max α
subject to

$$\text{Nes}(\bar{y}_i \subset \bar{a} \otimes \bar{x}_i) \geq \alpha, \quad i = 1, \dots, k, \\ (\bar{a}, c) \in L(J_1, J_2, J_3), \quad 0 < \alpha \leq 1.$$

$$\text{MP}_3: \max Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L), \quad (28)$$

max α
subject to

$$\text{Nes}(\bar{y}_i \supset \bar{a} \otimes \bar{x}_i) \geq \alpha, \quad i = 1, \dots, k, \\ (\bar{a}, c) \in L(J_1, J_2, J_3), \quad 0 < \alpha \leq 1.$$

3. Interactive algorithm for multiobjective programming problems

Since the fuzziness of the model and the fuzzy threshold α conflict with each other, the usual optimal solution concept for scalar optimization problems cannot be applied to the formulated multiobjective programming problems. Alternatively, for the multiobjective programming problems, the Pareto optimal solution concept is introduced, where any improvement of one objective function can be achieved only at the expense of another.

Definition 1. (\bar{a}^*, c^*) , α^* is said to be a Pareto optimal solution to the multiobjective programming problems MP_1 , MP_2 and MP_3 , if and only if there does not exist another (\bar{a}, c) , α such that $Z \leq Z^*$ for MP_1 and MP_2 ($Z \geq Z^*$ for MP_3), and $\alpha \geq \alpha^*$ with strict inequality holding for at least one objective function.

As can be immediately seen from Definition 1, usually, Pareto optimal solutions consist of an infinite number of points, and the decision maker must select his/her satisfying solution from among a Pareto optimal solution set based on his/her subjective judgement.

In the following, we make the following assumption.

Assumption 1. There implicitly exists an aggregation function $U(Z, \alpha)$ representing the degree of satisfaction of the decision maker for the fuzziness Z of the linear regression model and the fuzzy threshold α . This means that the decision maker can not specify the entire form of $U(z, \alpha)$, and he/she can provide local information concerning his/her preference. Moreover, $U(Z, \alpha)$ is strictly decreasing (increasing) with respect to Z for MP_1 and MP_2 (MP_3), and strictly increasing with respect to α .

Now in order to generate the Pareto optimal solutions to the multiobjective programming problems MP_1 , MP_2 and MP_3 , it is convenient to adopt the constraint problems [1] for some selected value ϵ ($0 < \epsilon \leq 1$).

In the following, we only focus on MP_1 , since similar results immediately follow for MP_2 and MP_3 .

The constraint problem for MP_1 is formulated as follows.

$$\text{CP}_1(\epsilon): \min Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L) \quad (29)$$

subject to

$$\text{Pos}(\bar{y}_i = \bar{a} \otimes \bar{x}_i) \geq \alpha, \quad i = 1, \dots, k, \\ (\bar{a}, c) \in L(J_1, J_2, J_3), \quad \alpha \geq \epsilon,$$

where ϵ is subjectively specified by the decision maker.

From Theorem 2, the constraint problem $\text{CP}_1(\epsilon)$ can be represented as the following linear programming problem with respect to (\bar{a}_j, c_j) , $j = 0, \dots, n$.

$$\text{CP}_1(\epsilon): \min Z = \sum_{i=1}^k (Y_{i0}^R - Y_{i0}^L) \quad (30)$$

subject to

$$\bar{y}_i + L^{-1}(\epsilon)e_i \geq Y_{ic}^L, \quad i = 1, \dots, k, \quad (31)$$

$$\bar{y}_i - L^{-1}(\epsilon)e_i \leq Y_{ic}^R, \quad i = 1, \dots, k, \quad (32)$$

$$\bar{a}_j - L^{-1}(\epsilon)c_j \geq 0, \quad j \in J_1, \quad (33)$$

$$\bar{a}_j - L^{-1}(\epsilon)c_j \leq 0, \quad \bar{a}_j + L^{-1}(\epsilon)c_j \geq 0, \quad j \in J_2, \quad (34)$$

$$\bar{a}_j + L^{-1}(\epsilon)c_j \leq 0, \quad j \in J_3, \quad (35)$$

$$c_j \geq 0, \quad j = 0, \dots, n \quad (36)$$

Note that the constraint problem $\text{CP}_1(\epsilon)$ is a linear programming problem with respect to (\bar{a}, c) .

Thus far, we have formulated the multiobjective linear programming problems for obtaining the fuzzy linear regression models under the assumption that the index set (J_1, J_2, J_3) satisfying the conditions (12), (13) and (14) is known in advance. However, even when the index set (J_1, J_2, J_3) satisfying the conditions (12), (13) and (14) is unknown in advance, we can also obtain the Pareto optimal solution to MP_1 by applying a linear programming based algorithm. For details see Algorithm 1 in the Appendix.

The relationships between the optimal solutions to $\text{CP}_1(\epsilon)$ and the Pareto optimal solutions to MP_1 can be characterized by the following theorem.

Theorem 3. (\bar{a}^*, c^*) and $\epsilon \in (0, 1]$ is a Pareto optimal solution to MP_1 , if and only if (\bar{a}^*, c^*) is a unique optimal solution to $\text{CP}_1(\epsilon)$ for $\epsilon \in (0, 1]$.

Now given a Pareto optimal solution for the degree ϵ specified by the decision maker for solving the constraint problem $\text{CP}_1(\epsilon)$, the decision maker must either be satisfied with the current Pareto optimal solution, or update the fuzzy threshold ϵ . In order to help the decision maker express his/her degree of preference, trade-off information between the fuzziness of the model and the fuzzy threshold ϵ is very useful. Such trade-off information is easily obtainable since it is closely related to the simplex multipliers of the constraint problems $\text{CP}_1(\epsilon)$. Having formulated the Lagrangian function L for $\text{CP}_1(\epsilon)$, and applying the sensitivity theorem [5], the following theorem can be obtained.

Theorem 4. For some fixed degree $\epsilon \in (0, 1]$, let (\bar{a}, c) be a unique and nondegenerate optimal solution to $\text{CP}_1(\epsilon)$, and let $\pi_i(\epsilon)$, $\lambda_i(\epsilon)$,

$i = 1, \dots, k$, and $\eta_j(\epsilon)$, $j = 0, \dots, n$, denote the simplex multipliers of the constraints (33), (34) and (34) respectively. Also let $Z(\epsilon)$ denote the optimal value functions to $\text{CP}_1(\epsilon)$. Then the following relations hold:

$$\frac{\partial Z(\epsilon)}{\partial \epsilon} = - \frac{\partial L^{-1}(\epsilon)}{\partial \epsilon} \sum_{i=1}^k \{ \pi_i e_i + \lambda_i e_i \} \\ + \sum_{i=1}^k \left\{ \pi_i \frac{\partial Y_{i0}^L}{\partial \epsilon} - \lambda_i \frac{\partial Y_{i0}^R}{\partial \epsilon} \right\} \\ + \frac{\partial L^{-1}(\epsilon)}{\partial \epsilon} \left\{ \sum_{j \in J_1} \eta_j c_j - \sum_{j \in J_2} \eta_j c_j + \sum_{j \in J_3} \eta_j c_j \right\}. \quad (37)$$

Following the above discussions, we can now present the linear programming based interactive algorithm in order to derive the satisfying solution for the decision maker from among the Pareto optimal solution set to MP_1 .

Interactive algorithm.

Step 1. Ask the decision maker to select the initial value of the fuzzy threshold ϵ ($0 < \epsilon \leq 1$).

Step 2. For the fuzzy threshold ϵ specified by the decision maker, solve the constraint problem $\text{CP}_1(\epsilon)$. (If the index set (J_1, J_2, J_3) satisfying the conditions (12), (13) and (14) is unknown, apply Algorithm 1 in the Appendix.)

Step 3. The decision maker is supplied with the Pareto optimal solution and the trade-off rate between the fuzziness of the model and the fuzzy threshold ϵ . If the decision maker is satisfied with the current values of the fuzziness and ϵ , stop. Otherwise, the decision maker must update the fuzzy threshold ϵ by considering the current values of the fuzziness and ϵ together with the trade-off rate between the fuzziness and ϵ , and return to Step 2.

The linear programming based interactive algorithm for obtaining the satisfying solution of the decision maker from among the Pareto optimal solution set to MP_2 and MP_3 follows very much in the same fashion as MP_1 and will, therefore, be omitted.

4. Numerical example

In order to demonstrate the appropriateness and the efficiency of the proposed method, we apply the proposed interactive algorithm to the eight pairs of fuzzy input-output data $(\tilde{x}_i, \tilde{y}_i)$, $i = 1, \dots, 8$, given in Table 1. In Table 1, x_i and y_i are the mean values of input-output data, d_i and e_i are the spread parameters. Assuming that the reference function $L(x) = 1 - |x|$, the membership functions of the fuzzy input-output data \tilde{x}_i, \tilde{y}_i are defined as follows:

$$\mu_{\tilde{x}_i}(u_i) = \begin{cases} 1 - |x_i - u_i|/d_i, & x_i - d_i \leq u_i \leq x_i + d_i, \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\tilde{y}_i}(v_i) = \begin{cases} 1 - |y_i - v_i|/e_i, & y_i - e_i \leq v_i \leq y_i + e_i, \\ 0, & \text{otherwise.} \end{cases}$$

In order to identify the fuzzy parameter vector $(\tilde{a}_0, \tilde{a}_1)$ of the fuzzy linear regression model,

$$\tilde{y}_i = \tilde{a}_0 + \tilde{a}_1 \otimes \tilde{x}_i,$$

the corresponding multiobjective programming problem MP_1 are formulated. The Pareto optimal solution to MP_1 is obtained by solving the constraint problem $CP_1(\epsilon)$ for some fixed degree ϵ . In this example, since it is known in advance that the any ϵ -level sets of the fuzzy parameters \tilde{a}_0 and \tilde{a}_1 are positive, we can directly obtain the optimal solution of $CP_1(\epsilon)$ without applying Algorithm 1 in the Appendix, where $J_1 = \{0, 1, \dots, n\}$, $J_2 = J_3 = \emptyset$. By applying the proposed interactive algorithm for MP_1 , the satisficing solution of the hypothetical decision maker is derived from among the

Table 1. Fuzzy input-output data

	Fuzzy input data		Fuzzy output data	
	x_i	d_i	y_i	e_i
1	2.0	0.5	4.0	0.5
2	3.5	0.5	5.5	0.5
3	5.5	1.0	7.5	1.0
4	7.0	0.5	6.5	0.5
5	8.5	0.5	8.5	0.5
6	10.5	1.0	8.0	1.0
7	11.0	0.5	10.5	0.5
8	12.5	0.5	9.5	0.5

Table 2. Interactive process

Iteration				
	1	2	3	4
ϵ	0.3	0.5	0.7	0.6
\tilde{a}_0	3.222	3.201	3.508	3.367
c_0	0	0.1703	0.9242	0.4260
\tilde{a}_1	0.5779	0.5788	0.5429	0.5595
c_1	0.04634	0.08059	0.1530	0.5595
Z	11.39	18.26	38.73	25.81
$\partial Z / \partial \epsilon$	20.72	60.43	178.8	94.23

Table 3. The optimal solution of $CP_1(\epsilon)$ and $CP_2(\epsilon)$ with $\epsilon = 0.6$

	$CP_1(\epsilon)$	$CP_2(\epsilon)$
objective function	25.81	58.34
\tilde{a}_0	3.367	3.996
c_0	0.4260	2.746
\tilde{a}_1	0.5595	0.4776
c_1	0.1108	0.07956

Pareto optimal solution set. In this example, at the fourth iteration, the satisficing solution of the hypothetical decision maker is obtained as shown in Table 2. Table 3 shows the comparison between the optimal solutions of $CP_1(\epsilon)$ and $CP_2(\epsilon)$ at the degree $\epsilon = 0.6$. Note that $CP_2(\epsilon)$ is infeasible for $\epsilon = 0.6$. In Figures 1 and 2, the eight rectangles are loci of the membership functions of the fuzzy input-output data $(\tilde{x}_i, \tilde{y}_i)$, $i = 0, \dots, 8$, whose values become 0. And the straight line and two broken lines are loci of the membership function of the fuzzy linear regression model $\tilde{a}_0 + \tilde{a}_1 \otimes \tilde{x}_i$, whose value becomes 1 and 0.6 respectively.

5. Conclusion

In this paper, using three indices for the equalities between two fuzzy numbers, we have formulated three types of multiobjective programming problems for obtaining fuzzy linear regression models where both input data and output data are fuzzy numbers. Then an interactive algorithm to derive the satisficing solution of the decision maker on the basis of the linear programming method was developed. Although a further detailed development would

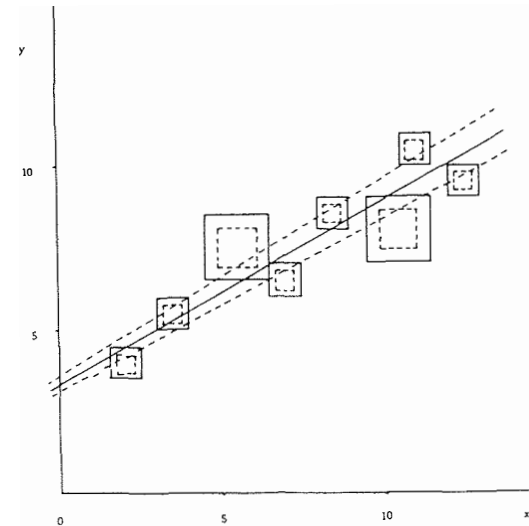


Fig. 1. The fuzzy linear regression model for MP_1 .

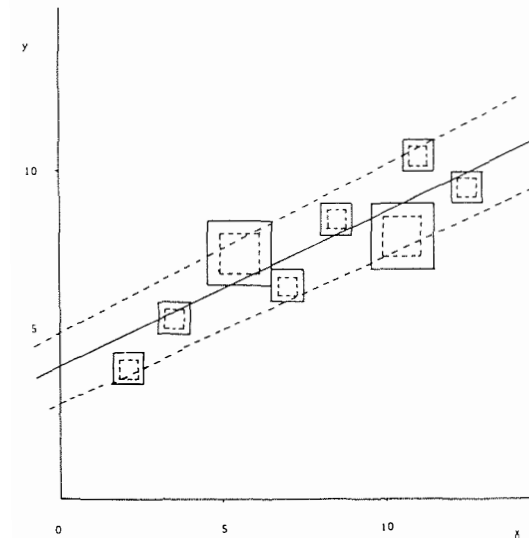


Fig. 2. The fuzzy linear regression model for MP_2 .

be necessary for fuzzy linear regression analysis, including the relationships to usual regression analysis as well as the application aspect, we hope that such fuzzy regression analysis would become efficient tools for analyzing real world systems where the ambiguity or fuzziness of human subjective judgement is influential.

Appendix

In the case where the index set (J_1, J_2, J_3) satisfying the conditions (12), (13) and (14) is unknown in advance, the Pareto optimal solution to MP_1 can be obtained as the accumulation point of the following linear programming based algorithm.

Algorithm 1.

Step 0. Set $J_1^{(1)} = \{0, 1, \dots, n\}$, $J_2^{(1)} = \emptyset$, $J_3^{(1)} = \emptyset$, and set the iteration index $l = 1$.

Step 1. Solve the following constraint problem to get the optimal solution $(\bar{a}_j^{(l)}, c_j^{(l)})$, $j = 0, \dots, n$:

$$CP_1(\epsilon; J_1^{(l)}, J_2^{(l)}, J_3^{(l)}): \min Z = \sum_{i=1}^k (Y_{i1}^R - Y_{i1}^L) \quad (A1)$$

subject to

$$\bar{y}_i + L^{-1}(\epsilon)e_i \geq Y_{i1}^L, \quad i = 1, \dots, k,$$

$$\bar{y}_i - L^{-1}(\epsilon)e_i \leq Y_{i1}^R, \quad i = 1, \dots, k,$$

$$\bar{a}_j - L^{-1}(\epsilon)c_j \geq 0, \quad j \in J_1^{(l)},$$

$$\bar{a}_j - L^{-1}(\epsilon)c_j \leq 0, \quad \bar{a}_j + L^{-1}(\epsilon)c_j \geq 0, \quad j \in J_2^{(l)},$$

$$\bar{a}_j + L^{-1}(\epsilon)c_j \leq 0, \quad j \in J_3^{(l)},$$

$$c_j \geq 0, \quad j = 0, \dots, n.$$

Step 2. If the optimal solution $(\bar{a}_j^{(l)}, c_j^{(l)})$, $j = 0, \dots, n$, of the constraint problem (A1) satisfies the following conditions, then stop:

$$\bar{a}_j^* - L^{-1}(\epsilon)c_j^* > 0, \quad j \in J_1^{(l)}, \quad (A2)$$

$$\bar{a}_j^* - L^{-1}(\epsilon)c_j^* < 0, \quad \bar{a}_j^* + L^{-1}(\epsilon)c_j^* > 0, \quad j \in J_2^{(l)}, \quad (A3)$$

$$\bar{a}_j^* + L^{-1}(\epsilon)c_j^* < 0, \quad j \in J_3^{(l)}, \quad (A4)$$

Otherwise, update the index set $J^{(l+1)}$ as follows:

$$J_1^{(l+1)} = \{j \in J_1^{(l)} \mid \bar{a}_j^{(l)} - L^{-1}(\epsilon)c_j^{(l)} > 0\} \\ + \{j \in J_2^{(l)} \mid \bar{a}_j^{(l)} - L^{-1}(\epsilon)c_j^{(l)} = 0\}, \quad (A5)$$

$$J_2^{(l+1)} = \{j \in J_2^{(l)} \mid \bar{a}_j^{(l)} - L^{-1}(\epsilon)c_j^{(l)} < 0 \\ \text{and } \bar{a}_j^{(l)} + L^{-1}(\epsilon)c_j^{(l)} > 0\} \\ + \{j \in J_1^{(l)} \mid \bar{a}_j^{(l)} - L^{-1}(\epsilon)c_j^{(l)} = 0\} \\ + \{j \in J_3^{(l)} \mid \bar{a}_j^{(l)} + L^{-1}(\epsilon)c_j^{(l)} = 0\}, \quad (A6)$$

$$J_3^{(l+1)} = \{j \in J_3^{(l)} \mid \bar{a}_j^{(l)} + L^{-1}(\epsilon)c_j^{(l)} < 0\} \\ + \{j \in J_2^{(l)} \mid \bar{a}_j^{(l)} + L^{-1}(\epsilon)c_j^{(l)} = 0\}. \quad (A7)$$

Set $l = l + 1$ and return to Step 1.

Concerning Algorithm 1, the following corollary holds.

Corollary 1. (1) Let $Z(\epsilon; J_1^{(l)}, J_2^{(l)}, J_3^{(l)})$ be an optimal function value of the constraint problem (A1). Then

$$Z(\epsilon; J_1^{(l)}, J_2^{(l)}, J_3^{(l)}) \geq Z(\epsilon; J_1^{(l+1)}, J_2^{(l+1)}, J_3^{(l+1)}), \quad (A8)$$

(2) If the conditions (A2), (A3) and (A4) are satisfied, then the optimal solution to (A1) is the local optimal solution to $CP_1(\epsilon)$.

(3) Any accumulation point of the optimal solution $(\bar{a}^{(l)}, c^{(l)})$ to (A1) is the local optimal solution of $CP_1(\epsilon)$.

Proof. The proofs of (1) and (2) are easy, and so omitted. In the following, we prove (3). Let (\bar{a}^*, c^*) be some accumulation point of the sequence $(\bar{a}^{(l)}, c^{(l)})$ in Algorithm 1. At the accumulation point (\bar{a}^*, c^*) , the index set $J = \{0, 1, \dots, n\}$ is divided as follows:

$$A_1 = \{j \in J \mid \bar{a}_j^* - L^{-1}(\epsilon)c_j^* > 0\},$$

$$A_2 = \{j \in J \mid \bar{a}_j^* - L^{-1}(\epsilon)c_j^* = 0\},$$

$$A_3 = \{j \in J \mid \bar{a}_j^* - L^{-1}(\epsilon)c_j^* < 0 \\ \text{and } \bar{a}_j^* + L^{-1}(\epsilon)c_j^* > 0\},$$

$$A_4 = \{j \in J \mid \bar{a}_j^* + L^{-1}(\epsilon)c_j^* = 0\},$$

$$A_5 = \{j \in J \mid \bar{a}_j^* + L^{-1}(\epsilon)c_j^* < 0\}.$$

There exists the following combinations of the index set J for $A_1 - A_5$:

$$(J_1, J_2, J_3) = (A_1 + A_2, A_3, A_4 + A_5),$$

$$(J_1, J_2, J_3) = (A_1 + A_2, A_3 + A_4, A_5),$$

$$(J_1, J_2, J_3) = (A_1, A_2 + A_3 + A_4, A_5),$$

$$(J_1, J_2, J_3) = (A_1, A_2 + A_3, A_4 + A_5).$$

Then it is clear that the index set of any element

of the sequence $(\bar{a}^{(l)}, c^{(l)})$ in the neighborhood $N(\bar{a}^*, c^*)$ is included among the above four types of combinations. Assume that (\bar{a}^*, c^*) is not a local optimal solution to $CP_1(\epsilon)$. Then there exists some (\bar{a}, c) in the neighborhood $N(\bar{a}^*, c^*)$ where the objective function value at (\bar{a}, c) give less than the objective function value at (\bar{a}^*, c^*) , and the corresponding index set (J_1, J_2, J_3) belongs to the above four types of combinations. This means that (\bar{a}^*, c^*) is not an optimal solution to (A1), where the index set (J_1, J_2, J_3) is any one of the above four types of combinations. \square

References

- [1] V. Chankong and Y.Y. Haimes, *Multiobjective Decision Making: Theory and Methodology* (North-Holland, Amsterdam, 1980).
- [2] D. Dubois, Linear programming with fuzzy data, in: J.C. Bezdek, Ed., *Analysis of Fuzzy Information*, Vol. 3, (CRC Press, Boca Raton, 1987) 241-263.
- [3] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Application* (Academic Press, New York, 1980).
- [4] D. Dubois and H. Prade, Ranking fuzzy numbers in the setting of possibility theory, *Inform. Sci.* **30** (1983) 183-224.
- [5] A.V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming* (Academic Press, New York, 1983).
- [6] H.T. Nguyen, A note on the extension principle for fuzzy sets, *J. Math. Anal. Appl.* **64** (1978) 369-380.
- [7] M. Sakawa and H. Yano, Fuzzy linear regression and its application to the sales forecasting, *Internat. J. on Policy and Inform.* **15** (1989) 111-125.
- [8] M. Sakawa and H. Yano, Multiobjective fuzzy linear regression analysis and its applications, *Trans. Inst. Electron. Inform. Comm. Engrs.* **J72-A** (5) (1989) 778-785 (in Japanese).
- [9] H. Tanaka, J. Watada and I. Hayashi, On three formulations of fuzzy linear regression analysis, *Trans. Soc. Instr. and Control Engrs.* **22** (1986) 1051-1057 (in Japanese).
- [10] H. Tanaka, Fuzzy data analysis by possibilistic linear models, *Fuzzy Sets and Systems* **24** (1987) 363-375.
- [11] H. Tanaka, I. Hayashi and J. Watada, Possibilistic linear regression analysis based on possibility measure, *Proceedings of Second IFSA Congress*, Tokyo (1987) 317-320.
- [12] L.A. Zadeh, The concept of linguistic variable and its application to approximate reasoning, *Inform. Sci.* **8** (1975) 199-249.

